Vol. 8, No. 2 June, 1954

Canadian Journal of Psychology

THE JOURNAL OF THE CANADIAN PSYCHOLOGICAL ASSOCIATION

CONTENTS

SPECIAL ISSUE: DEFENCE RESEARCH STUDIES

PUBLISHED QUARTERLY
MARCH - JUNE - SEPTEMBER - DECEMBER

THE UNIVERSITY OF TORONTO PRESS

$4.00 PER YEAR

Canadian Psychological Association

1953-1954

COUNCIL

Honorary President, W. LINE, *Toronto*

EXECUTIVE: President, D. C. WILLIAMS, Toronto; Past President, D. O. HEBB, Montreal; President Elect, FATHER NOËL MAILLOUX, Montreal; Secretary-Treasurer, G. A. FERGUSON, Montreal.

DIRECTORS: E. S. W. BELYEA, Vancouver; D. E. SMITH, Edmonton; R. B. MALMO, Montreal; B. M. SPRINGBETT, Montreal; J. M. BLACKBURN, Kingston; A. H. SMITH, Kingston.

STANDING COMMITTEES

MEMBERSHIP COMMITTEE: Chairman, DALBIR BINDRA, Montreal; F. T. SNODGRASS, Fredericton; LOUISE THOMPSON WELCH, Halifax; D. J. L. BÉLANGER, Montreal; G. DUFRESNE, Montreal; F. R. WAKE, Ottawa; JEAN BROWN, Toronto; A. H. SHEPHARD, Toronto, G. H. TURNER, London, G. A. MCMURRAY, Saskatoon; G. M. DUNLOP, Edmonton; E. I. SIGNORI, Vancouver; M. H. MUNRO, Vancouver.

COMMITTEE ON PUBLICATIONS: Chairman, D. C. WILLIAMS, Toronto; E. W. BOVARD, Toronto; A. H. SHEPHARD, Toronto; C. M. MOONEY, Toronto.

FINANCE COMMITTEE: Chairman, C. M. MOONEY, Toronto; with power to add.

COMMITTEE ON PROFESSIONAL STANDARDS: Chairman, O. E. AULT, Ottawa, E. A. BOTT, Toronto; G. A. FERGUSON, Montreal; L. H. ST. PIERRE, Montreal.

COMMITTEE ON SCIENTIFIC AND PROFESSIONAL ETHICS: Chairman, LOUISE THOMPSON WELCH, Halifax; S. N. F. CHANT, Vancouver; E. C. WEBSTER, Montreal; L. T. DAYHAW, Montreal; G. DUFRESNE, Montreal.

ELECTIONS COMMITTEE: Chairman, N. W. MORTON, Ottawa; with power to add.

C.P.A. REPRESENTATIVE ON THE CANADIAN SOCIAL SCIENCE RESEARCH COUNCIL: C. R. MYERS, Toronto.

1954 ANNUAL MEETING, June 5th and 6th

INTERNATIONAL CONGRESS OF PSYCHOLOGY, June 7th to 12th

MONTREAL, QUEBEC

CANADIAN JOURNAL OF PSYCHOLOGY: Correspondence regarding subscriptions should be sent to the Secretary-Treasurer, Canadian Psychological Association, 3600 McTavish St., Montreal, Que. Manuscripts or correspondence on editorial matters or advertising should be sent to the Editor, J. D. Ketchum, 100 St. George Street, Toronto 5, Ontario, Canada.

REGIONAL REPRESENTATIVES: E. S. W. Belyea (*British Columbia*); D. E. Smith (*Alberta*); G. A. McMurray (*Saskatchewan*); B. M. Springbett (*Manitoba*); J. M. Brown (*Ontario*); G. A. Ferguson, D. Bélanger (*Quebec*); W. H. D. Vernon (*Maritimes*).

The Canadian Psychological Association also issues THE CANADIAN PSYCHOLOGIST which is distributed to members only. EDITOR: Dr. Georges Dufresne, 49 Spring Grove Crescent, Montreal, P.Q., Canada.

AUTHORIZED AS SECOND-CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA.

Vol. 8, No. 2 June, 1954
Canadian Journal of Psychology

PSYCHOLOGY AND THE DEFENCE RESEARCH BOARD

N. W. MORTON
Defence Research Board, Ottawa

THIS issue of the *Canadian Journal of Psychology* provides a representative picture of the kind and quality of psychological and related research on defence problems which is being sponsored by the Defence Research Board. The papers come from members of the research staff of the Board, and from university colleagues whose work has been supported by the Board.

It is necessary to devote time and effort to the specific subject of national defence. We hope that this necessity may in time be at an end; none the less, it has revealed in the past ten or fifteen years certain things about psychology and psychologists which are of very general interest. It has demonstrated, in the first place, the practical relevance of psychological concepts and skills to such a vast undertaking as the mobilization of our armed forces. This, in turn has given confidence to those who may sometimes have wondered about the consistency or significance of the science, and has helped to generate a demand for more psychological research, both basic and applied. And, secondly, it has revealed the ability of psychologists, in and out of the services, to collaborate and co-operate in support of a common object.

The work reported here, and much other work which cannot be reported for reasons either of space or of military security, is the result of co-operative effort on the part of psychologists in the armed services, in the Defence Scientific Service, and in the universities. When the Defence Research Board was established in 1947, its constitution was founded on the premise of support by the scientific community of Canada. It has, indeed, served essentially as a trustee, in times of relative peace, for the entire corps of Canadian scientists who are called upon in time of total conflict to lend themselves to the national effort. Thus the programme of human resources research in the Defence Research Board has been guided to a great extent by its university advisers, who have also served as consultants on specific problems of the armed services. In turn, the Board has supplemented other sources of financial support for basic research in the universities, and for the training of graduate students, by making

grants for projects which seemed likely to be of defence interest. In the case of psychology and related fields, the value of these grants has amounted since 1947 to about a quarter of a million dollars, and the training of nearly a hundred students has been assisted.

Since 1950 an increasing research competence has also been developed within the armed services and the Board. This has been most evident in the size of staff and quality of facilities present in the Defence Research Medical Laboratories. Here, under the superintendence of Dr. M. G. Whillans, psychologists, sociologists, and kindred scientists, both in and out of uniform, have contributed, along with physiologists, biochemists, nutritionists, and others, to the solution of problems of the human factors in defence. Most of the problems of psychological interest have related ultimately to the simplification of the duties of the serviceman, the improvement of the equipment which he uses, the selection, classification, and training of military personnel, and the assessment of those social, physical, and biological features of the military environment which influence behaviour in it. Although much of this work has been in the field of applied psychology, there have been other researches of a more basic character; again, arrangements have been made in certain cases for basic studies in support of applied projects to be carried out in university laboratories interested in the problems involved.

It is hoped that readers of the *Canadian Journal of Psychology* will find this number, which has been produced by special arrangement and through the courtesy of the Editors, of interest and value.

A FACTORIAL STUDY OF CLOSURE[1]

C. M. MOONEY
Defence Research Medical Laboratories, Toronto

THE AIM of this study has been to develop some new tests that will facilitate the measurement of individual differences in the apprehension of incomplete or disorganized representations of commonplace things (closure).

The closure tests hitherto employed by Street (13), Leeper (5), Sheehan (12), Verville and Cameron (17), Verville (16), Thurstone (14, 15),Guilford and Lacey (3), Botzum (2), and Adkins and Lyerly (1), have not been extensive, and few have permitted simple group administration and scoring.

Accordingly, six purported closure tests, capable of simple group administration and scoring, have been constructed; these, with nine other mental tests, have been administered to some ninety young men; and the results have been factor-analysed to see if the tests define a closure dimension or dimensions.

THE CLOSURE CONCEPT

A conventional psychological definition of closure would be that given by Gardner Murphy (9):

> *Closure*: according to Gestalt theory, a basic principle whereby the tension initiated by a situation is resolved and the configuration (whether of behavior or of mental process) tends to as complete or "closed" a condition as the circumstances permit. An interrupted sneeze or a face in profile without a nose is an unclosed configuration which one tends to "complete."

The term was given psychological currency by Wertheimer, Koffka, and Köhler in their early expositions of Gestalt theory. It was one of the organizing forces—similarity, proximity, closure, good continuation—which determined the direction of perceptual organization. The early "crucial" illustrations of these factors were given in simple visual patterns. In these, closure has a literal aptness in describing the closing of gaps in lines, circles, triangles, and patterns.

Closure may be described in a more general sense as the moment of perceptual resolution; as the terminal phase of an act of perceptual contemplation; as the tension-relieving instant when meaning is ascribed to

[1]D. R. B. Project No. D 77-94-35-29.

or is recognized as emerging from a compelling constellation of objects or events. Its common manifestation would be the rapid, habitual recognition of commonplace objects and events—its measure being, presumably, in terms of perceptual speed. Beyond this, less frequent manifestations would be the comprehension of words, symbols, objects, sounds, and events which, though not completely represented, were yet sufficiently articulated to suggest their whole state. Least frequent manifestations would be those subjective and thematic perceptual constructions placed upon severely depleted or amorphous confrontations.

There would seem to be a conceptual kinship between insight and closure. Wertheimer (18) remarked, concerning insightful problem solving, that, for the person involved, "grasping the structure of the gap and the nature of the requirements that would enable him to close it adequately" was an instance of closure. "Often this transformation actually explodes, revolutionizes the old view" of the problem situation. Lewin (6) described "the act of insight" as "a reorganization of the field (Köhler) closely related in many respects to the transformation of so-called ambiguous figures . . . a shift in the totality of internal relationships." In the work of Hobhouse, Köhler, Yerkes, Duncker, Maier, and many others the general and consistent idea of insight was to the effect that it is the discovery of relationships—a kind of final precipitation of meaning, contingent on any or several of various predisposing conditions, occurring sometimes belatedly, sometimes early and suddenly, often complete and successful, sometimes not.

Should there be a practical distinction between closure and insight it would probably be this: insight may be long in coming, requiring close acquaintance with the problem, calculated essays at a solution, an accumulation of special knowledge, and a process of consolidation along the way leading to the final revelation of the answer. However, it may be more dramatic, coming as a sudden revelation not easily explained in retrospect, since it seems not to have hinged on any studied, rational approach, or, at least, when such was under way, to have been a sudden long leap forward to the answer. It would be this undeniable "clicking" or "aha!" experience which might be especially marked off as an instance of closure.

The clearest example is afforded by the simple instance of an incomplete picture of an everyday object, or a mutilated word. Here a specific object has been whittled away beyond immediate recognition; a few scattered bits and pieces, meaningless in themselves, together comprise a vestigial pattern or configural representation of a whole thing which an observer is to recognize or identify. He regards the configuration uncomprehendingly; as he contemplates it there presently occurs, if it occurs

at all, a sudden transformation (within the observer, obviously, since the configuration itself does not change) whereby the configuration seemingly snaps together and is then and thereafter, in some odd sense, "seen" as the whole object. It is a striking perceptual experience and gives rise to one of the classic problems in psychology—how it is that a present meaningless configuration can dictate the future emergence of an integrated structure by an appeal to organizing forces which must, presumably, stem from the unapprehended implications of past experience.

The Closure Tests

The postulated closure factor should reveal itself in a pronounced and uncomplicated manner in the apprehension of specific objects or ideas represented in incomplete or disorganized state. These representations should possess such "salient cues and reference points" (Cantril) as to create a compelling sense of imminent structure or wholeness. Recognition should come as a sudden and apparently fortuitous insight or closure, arising out of a process of contemplation rather than any species of rational analysis.

The following six closure tests were developed according to the foregoing prescription.

The *Closure Test* consists of incomplete and camouflaged drawings of familiar objects. There are 40 such items. The subject is to identify as many pictures as possible, in any order, in a period of 20 minutes. The reliability of this test, method of administration, scoring, and such details have been fully presented in an earlier paper by Mooney and Ferguson (8).

The *Mutilated Words Test* consists of 20 simple words presented in highly mutilated block letters. The words are to be identified in any order in a total period of 10 minutes.

The *Mutilated Sentences Test* contains 24 simple sentences spelled out in mutilated capital (block) letters. The sentences are to be deciphered and written out in a period of 10 minutes.

The *New Words Test* is made up of 70 nonsense words manufactured by prefixing the first letters of one word to the last letters of another word, where the two words are commonly used together. Thus SKRINK is to be identified as SKATING RINK. The time for this test is 10 minutes.

The *Spoonerisms Test* contains 30 simple sentences in which all the words have been spoonerized. Thus, LIVERS MOKE THE LOON is to be recognized as LOVERS LIKE THE MOON. The total time allowed is 10 minutes.

The *Disjointed Sentences Test* contains 20 simple sentences where the letters, in unchanged order, have been regrouped into nonsense words, with three or four letters omitted along the way. Thus: BEE RONTA INSAL COHOL is to be deciphered as: BEER CONTAINS ALCOHOL. Total time allowed is 10 minutes.

These six closure tests were incorporated into a battery with nine other tests—one being a verbal fluency test, four being presumed measures of some kind of rigidity, and four being measures of reasoning ability.

The *Word Fluency Test* was taken from the Chicago Tests of Primary Mental Abilities (L. L. Thurstone). It requires the subject to write down as many words as possible beginning with a given letter; and, again, to write down as many four-letter words as possible beginning with a particular letter. For this 9 minutes are allowed.

The *Alphabet Test* and *Arithmetic Test* were developed and used in the rigidity studies reported by Oliver and Ferguson (**10**) and by Scheier and Ferguson (**11**). The latter study also employed the *Reversed Reading Test*. These were found to define a non-motor or cognitive rigidity factor.

The *Alphabet Test* contains 52 items, each calling for that letter of the alphabet coming 2, 3, or 4 letters before a given letter. Thus M–3 calls for the response J. The time for this test is 5 minutes.

The *Arithmetic Test* consists of 60 simple arithmetic problems where a plus sign means subtract, a minus sign means add, a multiplication sign means divide, and a division sign means multiply. Time allowed is 5 minutes.

The *Reversed Reading Test* consists of 45 simple sentences with words in proper order but with the letters of each word in reverse order. The sentences are to be marked true or false. The time allowed is 5 minutes.

The *Hidden Words Test* represents an attempt to turn one of Luchins' (**7**) individual *Einstellung* or set tests into a group test. Items consist of several rows of letters; in each such row four or five letters, taken in order from left to right, are to be found which spell out a common word. Thus, in the row of letters PGETAXCNH there is hidden the word PEACH. In a given series the hidden words are all of a class—fruits, animals, colours, etc. The subject gets set to look for these. Later rows in the series contain intact four- or five-letter words which can be taken in place of the "hidden" class-words. The subject has to overcome the initially established set in order to see these. Six such items entailing discovery of some sixty words are to be dealt with in 9 minutes.

The *Following Directions Test* was suggested by Guilford (**3**). The subject is given a set of instructions to be applied along the way as he reads through a couple of pages of narrative. The injunctions are to underline certain kinds of words, circle other kinds, suspend a rule in favour of a substitute rule when any of the cue words are given in capital letters, and the like. The time allowed for this is 8 minutes.

The *Formations in Space Test* was also suggested by Guilford (**3**). It contains six items. An item is a verbal description of the related positions of three or four airplanes flying in formation. The formation is then to be plotted on 3 by 3 or 4 by 4 grids representing top, side, and oncoming views of the formation. The given flight formation has to be visualized and graphed in three-dimensional fashion. Total time for this is 10 minutes.

The remaining two tests are the well-known *Number Series* and *Verbal Analogies* sub-tests taken from Thurstone's A.C.E. Psychological Examination.

TABLE I

CORRELATION MATRIX AND FIFTH FACTOR RESIDUALS*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		−023	016	−029	−017	−015	−025	005	−039	045	027	042	007	013	−042
2	285		034	−020	002	008	016	−007	018	011	−042	−019	017	−019	−020
3	235	282		008	−075	−010	005	018	−018	−028	025	006	001	010	−043
4	160	137	285		039	009	024	−005	−045	044	−020	010	−006	−062	−008
5	388	228	337	392		028	023	048	−049	−026	009	009	−027	−061	029
6	248	166	517	490	625		005	−039	−006	−007	000	−022	−010	035	035
7	565	269	363	412	561	488		−012	020	−031	−038	−085	038	007	011
8	510	326	375	083	283	145	468		048	−041	−038	−028	−025	000	014
9	098	250	271	153	224	195	123	057		−020	039	020	044	−020	−029
10	145	215	292	310	394	373	293	054	166		−048	−015	−074	078	041
11	094	165	378	587	347	473	340	158	122	180		054	−005	−016	−003
12	−011	162	258	433	210	298	361	205	200	100	379		−014	000	−014
13	043	202	269	555	345	383	286	098	261	145	501	390		006	000
14	407	424	514	380	483	564	464	253	324	389	173	184	267		−041
15	338	203	157	320	302	191	349	058	134	334	145	149	278	424	

*Decimal points omitted. Residuals above principal diagonal, intercorrelation below.

The Testing Programme

The battery of tests was given to 92 Canadian Army Officer Training Corps cadets, undergraduates in engineering and the physical sciences. There were two group-testing sessions on successive evenings, taking about four hours in all. Testing conditions were ideal and the administration and timing of tests exactly controlled throughout.

The times allowed for the various tests were, in general, only sufficient for three or four persons to complete a test. The times allowed for the closure tests were derived from previous pilot trials. Times on Thurstone's tests were as prescribed.

Scoring was based on numbers of items answered correctly. These raw scores were normalized by tabulating their percentile ranks and assigning to each the corresponding standard score—thus giving normalized standard scores. These then formed the basis for calculating the correlations between all pairs of tests.

Analysis of the Results

The resulting correlation matrix, Table I, was factor-analysed by the Complete Centroid Method. Five factors were extracted. Fifth factor residuals were negligible and approximately normally distributed. These factor loadings were then rotated by Thurstone's Method of Extended Vectors to obtain the oblique simple structure revealed in Table II.

The oblique nature of this factor structure and the notable correlation revealed between Factors A, B, and C make for some difficulty in interpre-

TABLE II

Oblique Factor Matrix

	A	B	C	D	E
1	−007	−007	327	000	728
2	249	459	013	008	250
3	536	439	458	035	021
4	−006	−020	147	694	035
5	304	−015	514	234	221
6	452	001	680	274	−022
7	−001	010	374	355	585
8	016	385	216	−005	620
9	254	332	007	090	−017
10	356	078	245	070	013
11	−001	089	207	627	−006
12	−009	222	−012	473	043
13	016	133	003	604	−013
14	587	330	406	−024	175
15	018	002	−008	275	273

tation. Factors A, C, D, and E are the most compelling of the five factors. In the following interpretations major significance is attached to loadings of .300 and higher, and no significance to loadings below .200.

Factor A: Verbal Closure

14	Reversed Reading	.587
3	New Words	.536
6	Spoonerisms	.452
10	Word Fluency	.356
5	Disjointed Sentences	.304
9	Arithmetic Test	.254
2	Alphabet Test	.249

The top five tests call for facility in recognizing, thinking of, or trying out words that seem right or fit into a prescribed context. That the factor entailed is not simply or, perhaps, pre-eminently verbal fluency is suggested by the inferior position of the Word Fluency Test, and the peculiar characteristics of the tests ahead of it. The tests seem to demand quick, simple calling up or recognition of words that conform to an indicated prescription or pattern. This facility might be described as a simple instance of closure, and since the content is verbal in character, the factor might be described as *verbal closure.*

Factor B: Cognitive Rigidity

2	Alphabet Test	.459
3	New Words	.439
8	Closure Test	.385
9	Arithmetic Test	.332
14	Reversed Reading	.330
12	Verbal Analogies	.222

Notable here are the Alphabet Test, the Arithmetic Test, and the Reversed Reading Test. The tasks set by these are not simply resolvable by verbal facility; they call for procedures running counter to well-established habits. These three were definitive tests in the non-motor or cognitive rigidity factor identified by Oliver and Ferguson (**10**) and by Scheier and Ferguson (**11**). But it is somewhat surprising to find these associated with two of the purported closure tests—the New Words Test and the Closure Test. Of all the closure tests, these are, perhaps, most like riddles. In these a whole answer must come fortuitously to mind; in the quest, different answer possibilities must be called up for consideration until the one is found that fits the verbal or graphic specifications. It may be that the tests defining this present factor are of two kinds: those

that contain explicit specifications and those that contain implicit specifications; and one might surmise that the unifying axis is one having to do with speed or ability in calling up and selecting out conforming answers in the face of inhibiting or blocking features. Should this be so, performance on all of these tests might be expected to reflect any tendency towards what Duncker has called "functional fixedness," and what has already been called non-motor or cognitive rigidity. In describing this factor as *cognitive rigidity* one should keep in mind that it may be closely related to the factor that Thurstone (**14**) has called flexibility of closure.

Factor C: Conceptual Closure

6	Spoonerisms	.680
5	Disjointed Sentences	.514
3	New Words	.458
14	Reversed Reading	.406
7	Mutilated Sentences	.374
1	Mutilated Words	.327
10	Word Fluency	.245
8	Closure Test	.216
11	Formations in Space	.207

The six tests with significant loadings on this factor are all verbal in construction and five of them are purported closure tests. All six are capable of resolution by a sudden, confirming insight. Closure is contingent on a restructuring of the material in accordance with the conceptual or ideational relationship of the given parts to the concealed but clearly implied whole. The factor might, therefore, be appropriately designated *conceptual closure.*

Factor D: Formal Reasoning

4	Following Directions	.694
11	Formations in Space	.627
13	Number Series	.604
12	Verbal Analogies	.473
7	Mutilated Sentences	.355
15	Hidden Words	.275
6	Spoonerisms	.274
5	Disjointed Sentences	.234

The top four tests are of main interest here. The well-known Number Series and Verbal Analogies are typical measures of general reasoning. They call for logical operations of a categorical or syllogistic kind. The two most prominent tests, Following Directions and Formations in Space,

specify operations to be carried out according to prescribed rules where the rules and the operational data have to be recalled and logically applied under varying conditions. Examination of the last four tests suggests that they are in minor degree, at least, resolvable on logical-deductive grounds. It would seem in order, then, to describe Factor D as one of *formal reasoning*.

Factor E: Perceptual Closure

1	Mutilated Words	.728
8	Closure Test	.620
7	Mutilated Sentences	.585
15	Hidden Words	.273
2	Alphabet Test	.250
5	Disjointed Sentences	.221

This is evidently the perceptual closure factor revealed in studies by Thurstone (**14, 15**), Botzum (**2**), and Adkins and Lyerly (**1**). The three leading tests contain highly amorphous or depleted graphic presentations of familiar objects, letters, and words. The meaning of these presentations is determined by the structural congruence of the visible parts with the non-visible whole. Unlike the closure tests in Factor C, perception of the underlying meaning does not require any restructuring of the given materials. An act of completion is required whereby the enveloping whole is directly perceived.

Conclusions

Variations in scores made by many people on these 15 different tests can be accounted for in terms of five factors.

Three, and possibly four, factors entail closure—verbal (Factor A), conceptual (Factor C), and visual or perceptual (Factor E). A fourth factor, which has been called cognitive rigidity (Factor B) is closely related to these three closure factors and is evidently a factor of perceptual flexibility: non-flexibility. A fifth factor, formal reasoning (Factor D), is clearly marked off from the others. It refers to logical, deductive reasoning, whereas they refer to inductive or eductive—that is, insightful —perception.

Two of the non-closure tests created for this study, Following Directions and Formations in Space, are relatively pure and powerful measures of formal reasoning—reasoning which is categorically determined and is a kind of rote-reasoning, mechanical mentation, or mental "digitation"; what Guilford has described as "complex clerical-type operations" and Bartlett as "a serial sequence of inter-related moves."

Of the six closure tests, the Mutilated Words, Closure Tests, and Mutilated Sentences appear to be fairly substantial measures of strength of perceptual closure. The Spoonerisms and Disjointed Sentences tests reveal major variance on conceptual closure, with notable additional variance on verbal closure and some small variance on formal reasoning. The New Words test is factorially identical with Reversed Reading, both being compounded about equally of the verbal closure, cognitive rigidity, and perceptual closure factors. Thus, five of the six prospective tests created and put on trial—Spoonerisms, Disjointed Sentences, Closure Test, Mutilated Words, and Mutilated Sentences—commend themselves for individual development and standardization as effective measures of conceptual and perceptual closure.

REFERENCES

1. Adkins, C. D., & Lyerly, S. B. *Factor analysis of reasoning tests.* Personnel Res. Section Rep., No. 878. Washington, D.C.: Adjutant-General's Office.
2. Botzum, W. A. A factorial study of the reasoning and closure factors. *Psychometrika,* 1951, **16**, 361-386.
3. Guilford, J. P., & Lacey, J. T. (Eds.). *Printed classification tests.* Army Air Forces Aviation Psychol. Program Res. Rep., 1947, No. 5. Washington, D.C.: U.S. Government Printing Office.
4. Krech, D., & Crutchfield, R. S. *Theory and problems of social psychology.* New York: McGraw-Hill, 1948.
5. Leeper, R. A study of a neglected portion of the field of learning. *J. genet. Psychol.,* 1935, **46**, 41-75.
6. Lewin, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1936.
7. Luchins, A. S. *An examination for rigidity of behaviour.* Montrose, N.Y.: Veterans Administration Hospital, 1950.
8. Mooney, C. M., & Ferguson, G. A. A new closure test. *Canad. J. Psychol.,* 1951, **5**, 129–133.
9. Murphy, G. *Personality.* New York: Harper, 1947.
10. Oliver, J. A., & Ferguson, G. A. A factorial study of tests of rigidity. *Canad. J. Psychol.,* 1951, **5**, 49–59.
11. Scheier, I. H., & Ferguson, G. A. Further factorial studies of rigidity. *Canad. J. Psychol.,* 1952, **6**, 18–30.
12. Sheehan, M. R. A study of individual consistency in phenomenal constancy. *Arch. Psychol.,* 1938, 222.
13. Street, R. F. *A Gestalt completion test.* Teachers Coll., Columbia Univer. Theses, 1931, No. 4.
14. Thurstone, L. L. *A factorial study of perception.* Chicago: Univer. of Chicago Press, 1944.
15. Thurstone, L. L. *Mechanical aptitude III, analysis of group tests.* Univer. of Chicago Psychometr. Lab. Publ., 1949, No. 55.
16. Verville, E. The effect of emotional and motivational sets on the perception of incomplete figures. *J. genet. Psychol.,* 1946, **69**, 133–145.
17. Verville, E., & Cameron, N. Age and sex differences in the perception of incomplete pictures by adults. *J. genet. Psychol.,* 1946, **68**, 149–157.
18. Wertheimer, M. *Productive thinking.* New York: Harper, 1945.

A FACTORIAL STUDY OF INTEGRATION[1]

B. K. DOANE
Defence Research Medical Laboratories, Toronto

THE PURPOSE of this study was two-fold: first, to test the interpretations of three mental factors discovered during the Second World War, and second, to develop new group tests of complex mental operations.

During World War II, factor analysis was employed on a large scale in the development of new tests for aircrew selection in the United States Army Air Forces (1, 2). Among the many tests tried was a group known as "integration" tests. These were designed to measure a hypothetical ability to "integrate the influence of several simultaneously operative elements in a situation, all of which bear on the choice of a single direction of action" (2). The lack of such an ability, it was felt, might underlie many eliminations from flight training—eliminations ascribed to inabilities to perform a series of operations accurately or in proper order, to avoid repeating errors, or to exercise rapid decision, reaction, and judgment.

Seven tests designed to measure some such mental integration were combined in a battery with 24 other varied tests, administered to upwards of 250 aircrew trainees in the United States Army Air Forces, and the results factor-analysed. Two of the most important considerations in the construction of the seven so-called integration tests were: (1) that the cues to which the subject would respond should be presented in the midst of other activity and in such fashion as to require attention to a number of them simultaneously; and (2) that these cues should not lead to separate activities, but should require the integration or selection of a response or series of responses governed jointly by the separate cues.

When these tests were factor-analysed, 13 factors were found. Although the hypothetical integration factor did not appear, three of the 13 factors were seemingly new, looking unlike any which had been previously discovered; and these did seem to have some kinship with the initial integration concept. Accordingly, they were tentatively labelled Integration I, II, and III.

After the war these three factors received little attention, but later, when problems of aircrew selection again became important, it seemed worth while to study them further, particularly as some of the tests had shown validity for pilot selection. If reasonable hypotheses could be developed for integration factors I, II, and III, this might be a useful

[1]D. R. B. Project No. D 77–94–35–30.

starting point for further work with such abilities, whether or not the hypotheses needed subsequent revision. Accordingly, the U.S. wartime study was carefully reviewed, and the most plausible hypotheses for the three earlier factors were adopted as guides for constructing new tests which might be more suitable measures. Three tests were designed for each of the three hypotheses, and these were combined with two well-known reference tests of reasoning ability to form a battery of 11 tests.

Development of Hypotheses

The tests which had identified the factor called Integration I all presented the subject with a number of rules or injunctions at the outset. The essential requirements of the tests were seen before the test items were attacked, and these requirements remained unchanged throughout. The tests which distinguished Integration II differed in that new rules were added or existing ones modified during performance. The Integration III tests were more like those in Integration I, in that an initial set of rules remained fixed; however, whereas the rules were complex and the items simple in Integration I, in Integration III the rules were simpler and the items more complex.

Guilford (2) suggested that Integration I might be a memory factor, "the effective memory of a number of simple rules," or it might represent some functional aspect of mental set, "the effective persistence of a complicated mental set." Two of the most striking requirements for satisfactory performance on these tests were, however, speed and attention to details. Apparently, if given enough time, most subjects could solve the tests perfectly, but speed in shifting from one cue to the next might produce momentary lapses or breakdowns in performance, and the person would become confused. Accordingly, the appropriate operational definition for this factor seemed to be simply: *ability to shift frequently and rapidly from one to another of a succession of cues in applying a combination of simple rules.*

The first test constructed with this definition in mind was the Code Numbers Test, in which four simple, arbitrary, yet discrete rules had to be observed while the subject ran through successive columns of digits, underlining the appropriate number in each column depending upon which rule applied. The second was called the Serial Movements Test, a sample of which is shown in Figure 1. Here the subject had to count the spaces in a grid, keep track of the number of moves he had made, note the letters printed along the top of the grid, and repeatedly check the numbers printed in a row above and to the left of the grid. The third test in this group was the Transit Test. This differed somewhat from the other two in that not as many separate rules were given in the directions.

Here the subject merely had to keep track of the number of moves made by several imaginary ships diagrammed on the test sheet.

Guilford (2) spoke of Integration II as an "adaptability of mental set." He would have called it "flexibility of set (absence of perseveration)"

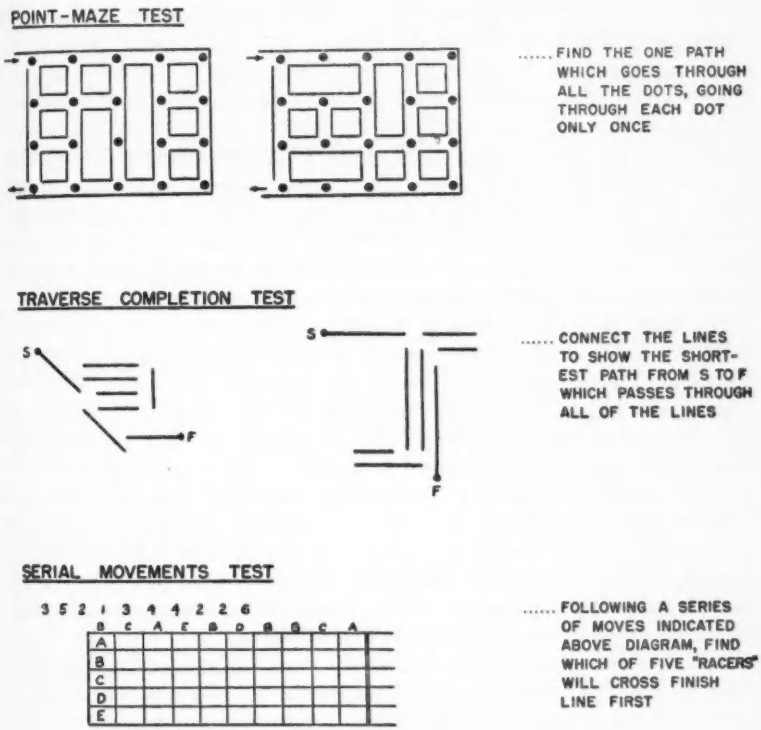

FIGURE 1. Samples of tests used in integration studies.

except for the fact that tests designed to measure flexibility or perseveration had repeatedly failed to show substantial intercorrelations. There was, however, an important difference between the familiar flexibility or perseveration tests and the Integration II tests. The former usually left the subject a choice between changing his response-set and continuing with the same one (3), whereas the Integration II tests offered no choice, but forced repeated changes upon him. In one the change was spontaneous, in the other enforced. Poor intercorrelations between the per-

severation tests could be attributed to the fact that they relied on subtle or hidden features which, since they were altogether apart from set, might elude the subject. The working definition adopted for this factor was: *ability to change or modify frequently a mental set when new rules or injunctions are added during the performance of a task.*

The first test in this group was the Following Directions Test, which closely resembled the test of the same name which had received the highest loading on the Integration II factor. Here the subject had to respond to a series of directions as he read them, being forced to alter his responses or add new considerations while in the process. The second test in this group was the Coins Test, in which the subject had to show how he would pile sets of coins, where the rules for piling them by age, value, and size were changed several times during the course of the test. The third was the Plotting Test, in which a series of points were to be plotted on a grid, calling for transposition from one graphic scale to another. The scales used were repeatedly changed throughout the test.

Guilford (**2**) offered three possible interpretations of Integration III. One was that it represented "span of apprehension or scope of apprehension." Another was that it might be a "mastery of details." A third, rather different from the other two, was that it represented "ideational fluency," or the "ease with which the individual can think of new possible responses." He thought the factor might best be summed up in the phrase "taking into account," though he was also impressed with the amount of planning apparently called for in these tests. In some ways it seemed to bear the closest resemblance to the original concept of integration, hence a reasonable hypothesis was that it involved *the ability to choose a proper response by means of rapid and thorough consideration of all the essential elements or cues in a situation.*

The first test designed for this group was the Traverse Completion Test, of which a sample is shown in Figure 1. This was essentially a maze-type test in which the subject was asked to join together a disjointed set of straight lines to indicate the shortest possible path (running through all of the lines) which would connect a starting point with a finishing point. The second was similar, and was called the Point Mazes Test; it is also illustrated in Figure 1. Here the subject was asked to draw a line showing the shortest possible path through a number of points in a maze. Last in this group was the Series Identification Test, which differed considerably from the other two. Here the subject had to select from five alternative series of numbers and letters that one which contained exactly the same numbers and letters as a key series, although the order within each series was changed. It seemed that this might group with the two

maze-type tests because several similar possible answers had to be checked to find the correct one.

The two reference tests were the Verbal Analogies and Number Series tests from Thurstone's A.C.E. Psychological Examination.

Administration and Scoring of the Tests

The eleven tests were administered to 111 flight cadets at the R.C.A.F. Station, London, Ontario. All subjects had senior matriculation or its equivalent. Testing was done in one evening, and took something over two hours in all, with a 15 minute rest period. The tests were timed to permit only a few subjects to complete any one test, and all scores were simply the number of correct answers. Raw scores were used in computing the intercorrelations, which are shown in Table I.

TABLE I

Correlation Matrix for the Eleven-Test Battery*

(N 111)

Test name and number	1	2	3	4	5	6	7	8	9	10	11
1. Code numbers	...	209	253	055	191	231	188	345	473	118	349
2. Serial Movements	...	...	238	246	290	223	292	259	237	076	241
3. Transit Test	...	...	...	096	459	418	425	391	176	282	522
4. Following Directions	...	...	...	...	558	409	181	228	117	503	346
5. Coins Test	...	...	...	...	...	529	269	347	216	481	492
6. Plotting Test	...	...	...	...	...	...	475	468	304	287	602
7. Traverse Completion	...	...	...	...	...	...	...	465	133	295	322
8. Point Mazes	...	...	...	...	...	...	...	...	274	268	236
9. Series Identification	...	...	...	...	...	...	...	...	...	217	433
10. Verbal Analogies	...	...	...	...	...	...	...	...	...	...	300
11. Number Series	...	...	...	...	...	...	...	...	...	...	...

*Decimal points omitted.

Factor Analysis

Thurstone's (7) complete centroid method was used in extracting the factors from the table of intercorrelations. Five factors were found. The fifth factor residuals were negligible and approximately normal in distribution, 48 being under .050 and the remainder under .110. The method of Extended Vectors (Thurstone) was used in rotation, which yielded a highly satisfactory oblique simple structure as shown in Table II.

Factor loadings less than .250 were considered to be of doubtful significance. Three of the five factors (Factors A, B, and C) are fair approximations to the hypotheses used in constructing the test battery. Factor A looks as follows:

Factor A	*Loading*
Code Numbers	.511
Series Identification	.443
Point Mazes	.423
Serial Movements	.292
Traverse Completion	.146
Verbal Analogies	.131

Factor A seems to come closest to meeting the hypothesis for Integration I. Two of the tests designed for that group are included here, namely, Code Numbers and Serial Movements. The Transit Test, also

TABLE II

ROTATED OBLIQUE FACTOR LOADINGS FOR THE ELEVEN-TEST BATTERY*

Test name and number	Factor				
	A	B	C	D	E
1. Code Numbers	511	−006	000	440	090
2. Serial Movements	292	003	008	000	295
3. Transit Test	031	004	007	−007	441
4. Following Directions	007	728	309	001	016
5. Coins Test	005	516	135	−013	330
6. Plotting Test	−035	372	369	290	007
7. Traverse Completion	146	−005	397	−014	101
8. Point Mazes	423	028	465	116	001
9. Series Identification	443	208	−003	572	002
10. Verbal Analogies	131	468	367	−003	004
11. Number Series	004	−388	−003	440	212

*Decimal points omitted.

intended for this group, is absent. Two of the tests designed for the definition of Integration III are also included here with significant loadings. A slight revision of the hypothesis seems necessary. This factor represents the shifts of attention, during maintenance of a complex mental set, which were sought for. The subject must not for a moment lose sight of all the things he is supposed to do while repeatedly shifting his attention from one thing to another. Precisely stated, this factor seems to represent the *ability to keep all elements of a complex mental set in effective operation throughout a task demanding attention to a succession of details.* It is difficult to choose a name for this factor. One might call it integration since it involves attention to numerous cues in following a prescribed course of action. It does not, however, lead to a single integrated response but, contrary to the original intent of the integration tests, calls for a series of separate activities. Such phrases as multiple

attention, successive attention, or serial operations might prove more fitting.

Factor B appeared as follows:

Factor B	*Loading*
Following Directions	.728
Coins Test	.516
Verbal Analogies	.466
Number Series	.388
Plotting Test	.372
Series Identification	.208

Factor B includes all three tests selected for our hypothesis regarding Integration II. The high loading of .728 received by the Following Directions Test indicates this as a strong element. At first glance it appears to support our interpretation of Integration II. One hesitates only because of the conspicuous presence of the two reference tests, Number Series and Verbal Analogies. These suggest a high involvement with reasoning ability. Because of the high loading of the distinctive Following Directions Test, one is inclined to accept the hypothesis that this is an *adaptability* of mental set, which might also be found in the two reference tests. One could argue, however, that this is speed of reasoning or speed of comprehension.

Loadings for Factor C were as follows:

Factor C	*Loading*
Point Mazes	.465
Traverse Completion	.397
Plotting Test	.369
Verbal Analogies	.367
Following Directions	.309
Coins Test	.135

Factor C is led by the two maze-type tests designed for Integration III. The third test designed for that group was of quite a different variety, and is missing from this list. Here, as in the case of Integration I, it is necessary to modify the hypothesis. This factor seems to represent the kind of "intelligence" measured in the well-known Porteus Mazes (4). It seems, therefore, to be a sort of "common sense" reasoning, or what might otherwise be called planning speed or planning capacity. This involves the *ability to plan and consummate an effective solution to a problem through the swift apprehension and organization of conditions and alternatives.*

Factor D appeared as follows:

Factor D	*Loading*
Series Identification	.575
Code Numbers	.440
Number Series	.440
Plotting Test	.290
Point Mazes	.116

Factor D is difficult to distinguish from Factor A, as the same two tests receive the highest loadings on both factors. Speed is a common feature among these tests, and the best interpretation seems to be that it is *speed of recognition*—the speed with which one can recognize something seen or described shortly beforehand. It might be the well-known number facility, but if so, it is surprising that the Number Series Test and Plotting Test, both of which call for actual arithmetical computations, did not receive the highest loadings.

Loadings for Factor E were as follows:

Factor E	*Loading*
Transit Test	.441
Coins Test	.330
Serial Movements	.295
Number Series	.212
Traverse Completion	.101
Code Numbers	.090

Factor E must be interpreted with caution, since only three of the tests have significant loadings, and these are not high. It may be a residual factor. However, there is a noticeable common feature among the tests, that of dealing with the relative order or positions of objects or of numbers. The factor might, then, represent a sense of order or arrangement, as with the letters in the alphabet, or in remembering a proper sequence of operations—a kind of *ordinal facility*, perhaps.

Concluding Observations

Conclusions drawn in this study are largely tentative. It has been necessary to qualify all three hypotheses for the factors in the U.S. Army Air Forces study. Were these tests to be administered with a wider variety of reference tests, some of the alternative interpretations might be eliminated. We seem to be dealing with a rather special type of mental abilities, which adds to the difficulty of naming Factors A and B.

The tests for Factor A lend themselves to the making of simple errors, apparently resulting from momentary breakdowns of fairly complex

operating sets, where a number of things must be attended to in a short span of time. The Code Numbers Test is the purest measure of this factor, and might be worth developing for the study of errors in routine, repetitive tasks.

Factor B looks like an important factor, though there is some doubt about its exact nature. One may assume, however, that it contributes to the kind of operational intelligence required in keeping up with ongoing activity in changing situations, as in aircraft or an air-control system. The Following Directions Test is decidedly worthy of development, its greatest lack being a simple method of scoring.

Factor C, which turned out to be a reasoning or planning factor, is probably the most fundamental of all, and the two maze-type tests might be worthy of further development.

While Factor D was hard to interpret, it may be the simplest factor of all, speed being its primary component. The Series Identification Test, which had the highest loading, has already been developed for experimental use in the Canadian Armed Forces.

Factor E was the most doubtful, yet the Transit Test was a pure measure of it in this study, and so might warrant further study, particularly if the hypothesis of ordinal facility is accepted.

Further development of any of these tests should be accompanied by a more comprehensive approach to the broad area of mental abilities with which we seem to be dealing, namely, those which function in complex, routine, repetitive activities. This should probably include psychophysical studies to ascertain the limits of central integrative processes under measured increments of speed and complexity.

REFERENCES

1. Flanagan, J. C. (Ed.). *The aviation psychology program in the army air forces.* Army Air Forces Aviation Psychol. Program, Res. Rep., 1948, No. 1. Washington, D.C.: U.S. Government Printing Office.
2. Guilford, J. P., & Lacey, J. I. (Eds.). *Printed classification tests.* Army Air Forces Aviation Psychol. Program, Res. Rep., 1947, No. 5. Washington, D.C.: U.S. Government Printing Office.
3. Luchins, A. S. *Examination for flexibility-rigidity.* Mimeographed rep. Montrose, N.Y.: Franklin D. Roosevelt Veterans' Administration Hospital.
4. Porteous, S. D. *The Porteus maze test and intelligence.* Palo Alto, Calif.: Pacific Books, 1950.
5. Spearman, C., & Jones, L. W. *Human ability.* London: Macmillan, 1950.
6. Thomson, G. H. *The factorial analysis of human ability.* (4th Ed.) London: Univer. of London Press, 1950.
7. Thurstone, L. L. *Multiple factor analysis.* Chicago: Univer. of Chicago Press, 1947.
8. Varney, A. *The psychology of flight.* Toronto: Van Nostrand, 1950.

EFFECTS OF DECREASED VARIATION IN THE SENSORY ENVIRONMENT[1]

W. H. BEXTON, W. HERON, AND T. H. SCOTT
McGill University

THIS STUDY began with a practical problem: the lapses of attention that may occur when a man must give close and prolonged attention to some aspect of an environment in which nothing is happening, or in which the changes are very regular. Watching a radar screen hour after hour is a prime example. As Mackworth (5) and others have shown, when at last something *does* happen in such circumstances the watcher may fail to respond. Such monotonous conditions exist in civilian occupations as well as in military ones (marine pilotage by radar, piloting aircraft on long flights), and here too lapses of attention may have extremely serious consequences. For example, such lapses may explain some otherwise inexplicable railroad and highway accidents.

Besides its practical significance this problem has theoretical implications of great interest. There is much evidence from recent neurophysiological studies to indicate that the normal functioning of the waking brain depends on its being constantly exposed to sensory bombardment, which produces a continuing "arousal reaction." Work now being done by S. K. Sharpless at McGill indicates, further, that when stimulation does not change it rapidly loses its power to cause the arousal reaction. Thus, although one function of a stimulus is to evoke or guide a specific bit of behaviour, it also has a non-specific function, that of maintaining "arousal," probably through the brain-stem reticular formation.

In other words, the maintenance of normal, intelligent, adaptive behaviour probably requires a continually varied sensory input. The brain is not like a calculating machine operated by an electric motor which is able to respond at once to specific cues after lying idle indefinitely. Instead it is like one that must be kept warmed up and working. It seemed, therefore, worth while to examine cognitive functioning during prolonged perceptual isolation, as far as this was practicable. Bremer (2) has achieved such isolation by cutting the brain stem; college students, however, are reluctant to undergo brain operations for experimental purposes, so we had to be satisfied with less extreme isolation from the environment.

[1]D. R. B. Project No. D 77-94-85-01. The advice and assistance of Dr. D. O. Hebb are gratefully acknowledged.

Procedure

The subjects, 22 male college students, were paid to lie on a comfortable bed in a lighted cubicle 24 hours a day, with time out for eating and going to the toilet. During the whole experimental period they wore translucent goggles which transmitted diffuse light but prevented pattern vision. Except when eating or at the toilet, the subject wore gloves and cardboard cuffs, the latter extending from below the elbow to beyond the fingertips. These permitted free joint movement but limited tactual perception. Communication between subject and experimenters was provided by a small speaker system, and was kept to a minimum. Auditory stimulation was limited by the partially sound-proof cubicle and by a U-shaped foam-rubber pillow in which the subject kept his head while in the cubicle. Moreover, the continuous hum provided by fans, air-conditioner, and the amplifier leading to earphones in the pillow produced fairly efficient masking noise.

General Effects

As might be expected from the evidence reviewed by Kleitman (3) for onset of sleep following reduced stimulation in man and other animals, the subjects tended to spend the earlier part of the experimental session in sleep. Later they slept less, became bored, and appeared eager for stimulation. They would sing, whistle, talk to themselves, tap the cuffs together, or explore the cubicle with them. This boredom seemed to be partly due to deterioration in the capacity to think systematically and productively—an effect described below. The subjects also became very restless, displaying constant random movement, and they described the restlessness as unpleasant. Hence it was difficult to keep subjects for more than two or three days, despite the fact that the pay ($20 for a 24-hour day) was more than double what they could normally earn. Some subjects, in fact, left before testing could be completed.

There seemed to be unusual emotional lability during the experimental period. When doing tests, for instance, the subjects would seem very pleased when they did well, and upset if they had difficulty. They commented more freely about test items than when they were tested outside. While many reported that they felt elated during the first part of their stay in the cubicle, there was a marked increase in irritability toward the end of the experimental period.

On coming out of the cubicle after the experimental session, when goggles, cuffs, and gloves had been removed, the subjects seemed at first dazed. There also appeared to be some disturbance in visual perception, usually lasting no longer than one or two minutes. Subjects reported difficulty in focussing; objects appeared fuzzy and did not

stand out from their backgrounds. There was a tendency for the environment to appear two-dimensional and colours seemed more saturated than usual. The subjects also reported feelings of confusion, headaches, a mild nausea, and fatigue; these conditions persisted in some cases for 24 hours after the session.

Effects on Cognitive Processes

Our present concern is primarily with cognitive disturbances during the period of isolation and immediately afterwards. The subjects reported that they were unable to concentrate on any topic for long while in the cubicle. Those who tried to review their studies or solve self-initiated intellectual problems found it difficult to do so. As a result they lapsed into day-dreaming, abandoned attempts at organized thinking, and let their thoughts wander. There were also reports of "blank periods," during which they seemed unable to think of anything at all.

In an attempt to measure some of the effects on cognitive processes, various tests were given to the subjects before, during, and after the period of isolation.

First, the tests given during isolation. Twelve subjects were given the following types of problem to do in their heads: multiplying two- and three-digit numbers; arithmetical problems (such as "how many times greater is twice 2⅛ than one-half 2½?"); completion of number series; making a word from jumbled letters; making as many words as possible from the letters of a given word. Each subject was tested on problems of this type before going into the cubicle, after he had been in for 12, 24, and 48 hours, and three days after coming out of the cubicle. Twelve control subjects were given the same series of tasks at the same intervals. The average performance of the experimental subjects was inferior to that of the controls on all tests performed during the cubicle session. With our present small number of subjects the differences are significant only for the error scores on the second anagram task ($p = .01$, see Figure 1). The groups are now being enlarged.

Secondly, tests given before entering the cubicle and immediately after leaving it. On the Kohs Block Test and the Wechsler Digit Symbol Test the experimental subjects were inferior to the controls on leaving the cubicle ($p = .01$). They also tended to be slower in copying a prose paragraph ($p = .10$). Figure 2 gives samples of handwriting before and after the experiment. The first is from one of the subjects showing the greatest effect, the second illustrates the average effect. As the third sample shows, some subjects were not affected. This disturbance in handwriting, though perhaps due to some sensori-motor disturbance, might also reflect cognitive or motivational changes.

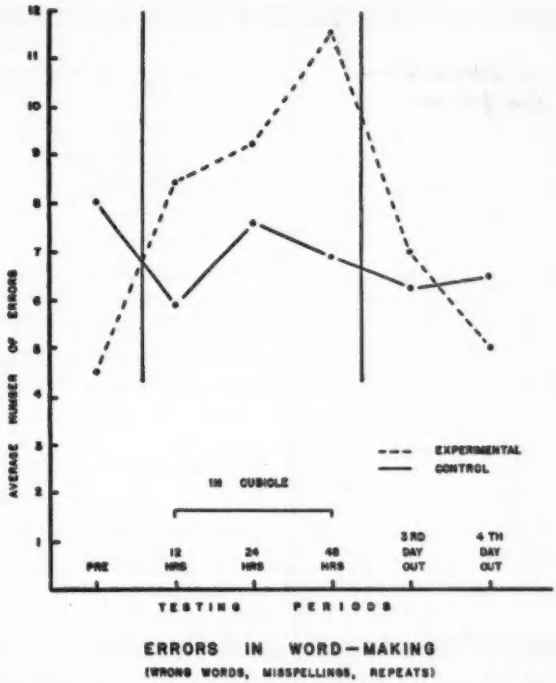

FIGURE 1. Mean error scores for experimental and control subjects, before, during, and after the isolation period.

HALLUCINATORY ACTIVITY

Finally there were the hallucinations reported by the subjects while in the experimental apparatus. Among our early subjects there were several references, rather puzzling at first, to what one of them called "having a dream while awake." Then one of us, while serving as a subject, observed the phenomenon and realized its peculiarity and extent.

The visual phenomena were actually quite similar to what have been described for mescal intoxication, and to what Grey Walter (6) has recently produced by exposure to flickering light. There have also been rare cases of hallucinations in aged persons without psychosis (1), which, like ours, involved no special chemical or visual stimulation. As we did not ask our first subjects specifically about these phenomena we do not know the frequency among them. The last 14 subjects, however, were asked to report any "visual imagery" they observed, and our report is

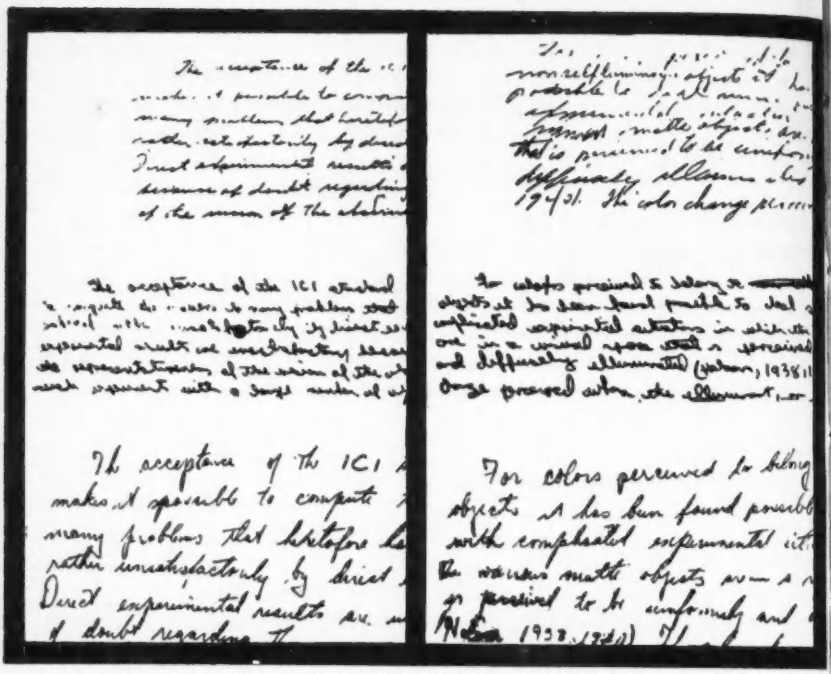

FIGURE 2. Specimens of handwriting before and after the isolation period.

based on them. In general, where more "formed" (i.e., more complex) hallucinations occurred they were usually preceded by simpler forms of the phenomenon. Levels of complexity could be differentiated as follows: In the simplest form the visual field, with the eyes closed, changed from dark to light colour; next in complexity were dots of light, lines, or simple geometrical patterns. All 14 subjects reported such imagery, and said it was a new experience to them. Still more complex forms consisted in "wall-paper patterns," reported by 11 subjects, and isolated figures or objects, without background (e.g., a row of little yellow men with black caps on and their mouths open; a German helmet), reported by seven subjects. Finally, there were integrated scenes (e.g., a procession of squirrels with sacks over their shoulders marching "purposefully" across a snow field and out of the field of "vision"; prehistoric animals walking about in a jungle). Three of the 14 subjects reported such scenes, frequently including dreamlike distortions, with the figures often being described as "like cartoons." One curious fact is that some of the hallucinations were reported as being inverted or tilted at an angle.

In general, the subjects were first surprised by these phenomena, and then amused or interested, waiting for what they would see next. Later, some subjects found them irritating, and complained that their vividness interfered with sleep. There was some control over content; by "trying," the subject might see certain objects suggested by the experimenter, but not always as he intended. Thus one subject, trying to "get" a pen, saw first an inkblot, then a pencil, a green horse, and finally a pen; trying to "get" a shoe, he saw first a ski boot, then a moccasin. The imagery usually disappeared when the subject was doing a complex task, such as multiplying three-place numbers in his head, but not if he did physical exercises, or talked to the experimenter.

There were also reports of hallucinations involving other senses. One subject could hear the people speaking in his visual hallucinations, and another repeatedly heard the playing of a music box. Four subjects described kinesthetic and somesthetic phenomena. One reported seeing

FIGURE 3. Drawing made by a subject to show how he felt at one period in the cubicle. He reported that it was as if "there were two of me," and was momentarily unable to decide whether he was A or B.

a miniature rocket ship discharging pellets that kept striking his arm, and one reported reaching out to touch a doorknob he saw before him and feeling an electric shock. The other two subjects reported a phenomenon which they found difficult to describe. They said it was as if there were two bodies side by side in the cubicle; in one case the two bodies overlapped, partly occupying the same space. Figure 8 shows this subject's subsequent drawing, made in an attempt to show what he meant.

In addition, there were reports of feelings of "otherness" and bodily "strangeness" in which it was hard to know exactly what the subject meant. One subject said "my mind seemed to be a ball of cotton-wool floating above my body"; another reported that his head felt detached from his body. These are familiar phenomena in certain cases of migraine, as described recently by Lippman (4), and earlier by Lewis Carroll in *Alice in Wonderland*. As Lippman points out, Lewis Carroll was a sufferer from migraine, and it is suggested that Alice's bodily distortions are actually descriptions of Carroll's (i.e., Charles Dodgson's) own experiences.

In summary, both the changes in intelligence-test performance and the hallucinatory activity, induced merely by limiting the variability of sensory input, provide direct evidence of a kind of dependence on the environment that has not been previously recognized. Further experimental study will be needed to elucidate the details of this relationship.

REFERENCES

1. Bartlet, J. E. A. A case of organized visual hallucinations in an old man with cataract and their relation to the phenomena of the phantom limb. *Brain*, 1951, 74, 363–373.
2. Bremer, F. & Terzuolo, C. Nouvelles recherches sur le processus physiologique de réveil. *Arch. internat. de Physiol.*, 1953, 61, 86–90.
3. Kleitman, N. *Sleep and wakefulness*. Chicago: Univer. of Chicago Press, 1939.
4. Lippman, Caro. Certain hallucinations peculiar to migraine. *J. nerv. ment. Dis.*, 1952, 116, 346–351.
5. Mackworth, N. H. *Researches on the measurement of human performance*. Med. Res. Council, Spec. Rep. Ser., 1950, No. 268. London.
6. Walter, W. Grey. *The living brain*. New York: Norton, 1953.

EFFECT OF FORM ON THE LEGIBILITY OF NUMBERS[1]

H. LANSDELL

Defence Research Medical Laboratories

THE LITERATURE on legibility of numbers consists largely of reports on the effects of such variables as stroke width, level of illumination, type of illumination, height-width ratio of the numbers, spacing, and colour contrast. That there has not been systematic research on number form can be excused because of the difficulty of defining possible types of forms. At present, reference is generally made to three types of numbers (4). One set is that recommended by the American Aeronautical Board, another was designed by Berger, and the third by Mackworth (2). The last two types have more straight strokes, and more sharp ends to the strokes, than the first. The Mackworth digits are probably the best (6). The emphasis on sharpness in stroke is found in other attempts at redesign of numbers (1). The present experiment may be thought of as carrying this trend towards angularity further. Since the triangle and square are two of the more "discriminable" forms (3), an attempt was made to use angular forms to represent numbers, breaking away from the use of a standard stroke width. The principle was that of using triangular and rectangular components to design forms as nearly as possible like "normal" numbers.

PROCEDURE

The legibility of the new angular forms was compared with that of Mackworth's numbers and a revision of the latter by Mound (5) in which he used varying stroke and number width to increase the legibility (Figure 1).

Each of the three types of numbers was shown at a distance of 14 feet to six observers. The digits were black on a white background and ¼" high. One digit at a time appeared inside and at the back of a dark box, 2′ x 2′ x 3′ long; an opening in the front of the box permitted the figure to be seen from *O*'s seat. The digit was illuminated for 0.6 seconds by a shaded photoflood bulb inside the box, which gave a brightness of about 10 footlamberts. The exposure time was controlled by an electronic timer. The digit was revealed when *O* activated the timer by pressing a button. This button would operate only when a small signal light on the front of the apparatus was lighted; the signal light was controlled by the experimenter with a foot switch. This ensured that *O* did not push the button while a number was being changed. The room was illuminated by indirect lighting from overhead fluorescent tubes.

[1]D. R. B. Project No. D 77-94-20-21.

Each *O* was given an instruction sheet explaining the procedure, and a photograph of the three types of numbers. A practice session of 15 or 20 minutes preceded the test. During the test each *O* was shown 600 numbers in all, 200 of each of the three types. One hundred numbers of one type were presented, then 100 of a second type, followed by 100 of the third type; then the order was reversed and three groups of 100 presented again. The *O*s wrote down the numbers, as they were shown, on a record sheet. They pushed the button at a rate of about once every 1½ seconds. When not sure of a number, the *O*s were required to guess. The numbers were presented according to a list in which each of the ten digits appeared ten times in various permutations.

FIGURE 1. The new angular forms, the Mound, and the Mackworth numbers.

RESULTS

No difference in legibility was found between the Mackworth numbers and the Mound revision of them. The overall percentage correct for the Mackworth numbers was 51.5, for the revision, 51.4. The percentage correct for the new angular forms was 67.4. All six observers did better on the new angular forms, which gives a .03 significance level with the Friedman chi-square test of rank (7).

Discussion

The better legibility of the new angular forms may be related to the fact that they were presented singly. A preliminary study of them in groups of five digits did not show the same superiority. Mound's recent revision of the Mackworth numbers was based on research with such groups of numbers. It may well be that under normal conditions the angular forms in groups of five digits are not more efficient than the Mound revision, and that both these forms are only slightly better than the Mackworth numbers.

The difficult viewing conditions may also have been partly responsible for the better legibility of the new forms. These conditions were imposed in order to reduce the number of observations necessary to obtain data on confusion between specific numbers. Such data are needed for further work on the design.

Summary

A new set of numbers was designed which breaks away from the tradition of using a uniform stroke width for number forms, and instead makes as much use as possible of more easily discriminated form components. With little practice the numbers proved to be better than an accepted standard set, when presented singly under difficult viewing conditions.

REFERENCES

1. Atkinson, W. H., Crumley, L. M., & Willis, M. P. A study of the requirements for letters, numbers, and markings to be used on trans-illuminated aircraft control panels. *U.S. Naval Air Material Center Rep. Ted.*, 1952, No. NAM E1–609, pt. 5.
2. Bartlett, F., & Mackworth, N. H. *Planned seeing*. London: H. M. Stationery Office, 1950.
3. Casperson, R. C. The visual discrimination of geometric forms. *J. exp. Psychol.*, 1950, **40**, 668–681.
4. Fitts, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
5. Mound, H. Personal communication.
6. Schapiro, H. B. Factors affecting legibility of digits. *USAF, Wright Air Development Center, Tech. Rep.*, 1952, No. 52–127.
7. Wilcoxon, F. *Some rapid approximate statistical procedures*. New York: American Cyanimid Co., 1949.

NON-VISUAL PERCEPTION OF THE POSTURAL VERTICAL: II. LATERAL PLANE[1]

W. C. CLEGG AND N. M. DUNFIELD[2]

Defence Research Medical Laboratories, Toronto, Ont.

A PREVIOUS REPORT (1) dealt with the effects of four variables upon the perception by male subjects of the postural upright following body tilt in the sagittal plane. This is a report of four experiments which tested the effects of the same variables in the lateral plane upon male and female subjects. The variables studied were: (*a*) the degree of initial tilt, i.e. the degree to which the subject was tilted before being moved toward the vertical, (*b*) the time held at initial tilt, (*c*) the direction of initial tilt, and (*d*) the angular velocity of return.

Other studies in this area have been summarized in the symposium on *Psychophysiological Factors in Spatial Orientation* (7), particularly in the articles by Gibson, Mann, and Witkin. It has been reported that the error of judgment of the vertical increases in the direction of initial tilt as the magnitude of initial tilt increases (1, 2, 4, 5) and as the velocity of return decreases (1, 2). Studies of the effect of direction of tilt within any one plane are not in agreement. One investigation (6) found no effect upon error of judgment attributable to direction, whereas others (1, 2) revealed significant differences in judgment between the two directions in a single plane. Witkin (7) reported sex differences in the perception of the vertical, females being affected more than males by visual cues when they were in conflict with postural cues. Time held at initial tilt was found to influence the error of judgment in certain studies (3, 4, 5, 6), although other studies (1, 8) revealed no relationship between duration of tilt and error of judgment of the upright.

APPARATUS

The apparatus has been described in detail elsewhere (1). It consisted of a chair in which blindfolded subjects, with their heads fixed, could be tilted to various positions right or left of the gravitational vertical and returned at various velocities. A large protractor, readable to within 1/8°, indicated the degree of tilt of the apparatus.

GENERAL PROCEDURE

The general procedure was identical with that reported earlier (1).

[1]D. R. B. Project No. D 77-94-35-36.

[2]At present with Imperial Oil Limited, Department of Employee Relations.

Subjects were tilted to various positions right or left of the vertical, held there for varying intervals of time, and returned toward the vertical at various speeds. When they judged that they were upright, they responded by sounding a buzzer.

The order of presentation of the variables was random and each subject was tested once for each combination of the variables.

The experiments differed in that in Experiment I direction and magnitude of tilt, time at initial tilt, and velocity of return were varied; in Experiment II velocity, direction, and magnitude of tilt were held constant and time alone varied; in Experiment III velocity was variable with the other three constant. All subjects in Experiments I, II, and III were male. Experiment IV was a repetition of Experiment I using female subjects.

Experiment I

Two directions (left and right), three angles (9°, 12°, 15°), three delay times (15, 30, 45 seconds), and three velocities (6.5°, 14.0°, and 17.5° per minute) were presented in a $2 \times 3 \times 3 \times 3$ factorial design.

Results and Discussion

An analysis of variance, Table I, revealed that there are significant differences in degree of mean error attributable to subjects, direction of initial tilt, speed of return, and magnitude of initial tilt. All these differences are significant beyond the 1 per cent level. Table I reveals no significant effect of delay time upon the magnitude and direction of error.

Table II shows the mean errors obtained with the different values of the experimental variables. It can be seen that the mean error is smaller when subjects return from tilt to the left than from right tilt; that the mean error decreases as speed of return increases; and that the mean error is an increasingly negative value as the magnitude of initial tilt increases.

The significant difference in mean error between the two directions of tilt supports the findings reported by Fleishman (**2**) and essentially parallels those reported by Clegg and Dunfield (**1**) respecting sagittal tilt. In the sagittal tilt experiments it was suspected that the apparatus might have been responsible for the differences in error ascribable to the two directions. To tilt subjects laterally, the apparatus was altered in such a manner that artifacts within its structure presumed to cause larger errors from backward tilt would have given rise to larger errors from left tilt. That this did not occur rules out the possibility that the differences for direction are apparatus artifacts. Fleishman's (**2**) experiments used an

TABLE I

ANALYSIS OF VARIANCE TABLE
2 × 3 × 3 × 3 FACTORIAL WITH 10 REPLICATIONS (S's)

Source of variation	d.f.	Sums of squares	Mean squares	F-ratio
Subjects	9	795.50	88.39	16.40**
Direction (Q)	1	361.40	361.40	67.05**
Delay time (T)	2	9.52	4.67	0.88
Speed of return (S)	2	266.75	133.38	24.75**
Degree of tilt (D)	2	244.71	122.36	22.70**
Q × T	2	3.95	1.98	0.37
Q × S	2	2.23	1.12	0.21
Q × D	2	0.83	0.42	0.08
T × S	4	16.76	4.19	0.78
T × D	4	11.31	2.83	0.52
S × D	4	22.76	5.69	1.06
Q × T × S	4			
Q × T × D	4			
Q × S × D	4			
T × S × D	8			
Q × T × S × D	8			
Error	477	2569.09	5.39	
Total	539	4304.81		

**Significant at the 1% level.

TABLE II

MEAN ERRORS FOR EACH VARIABLE

Variable		$\bar{X}$
Direction	Right	−2.04
	Left	−0.41
Return speed	6.5°/min.	−2.13
	14.0°/min.	−1.14
	17.5°/min.	−0.42
Degree of tilt	9°	−0.42
	12°	−1.21
	15°	−2.06
Delay time	15 secs.	−1.08
	30 secs.	−1.21
	45 secs.	−1.40

entirely different type of apparatus, yet similar results were obtained. No adequate explanation of this phenomenon is as yet available.

The observations concerning mean error as influenced by "degree of initial tilt" are in general agreement with previous studies (1, 2, 3, 4, 5, 6).

"Delay time" showed no significant effect on error. This finding agrees with the earlier experiments in the sagittal plane (1) but not with those in the lateral plane (3, 4, 5, 6). In the hope of resolving this disagreement, Experiment II was conducted.

The effect of "velocity of return" was the same as that found in the sagittal plane experiments. Experiment III was conducted to define this relationship more fully.

Experiment II

Subjects were tilted to the right to a position of 15.0° at a speed of 12.0° per minute and were held there for 0, 5, 10, 15, 30, 45, 60, 75, 90, and 105 seconds before returning toward vertical at a speed of 17.5° per minute.

Results and Discussion

An analysis of variance revealed significant differences in degree of error attributable to subjects, but no relationship between duration of tilt and the mean error of judgment (Table III).

TABLE III

Analysis of Variance Table

Source	d.f.	Sums of squares	Mean squares	F-ratio
Subjects	9	144.6257	16.0695	3.39**
Time tilted	9	35.3257	3.9251	0.82
Error	81	389.8805	4.8133	
Total	99	569.8319		

**Significant beyond the 1% level of confidence.

Table IV presents the means and S.D. of error for the various times as well as the frequency of overestimations, underestimations, or correct estimates. It can be seen that mean error does not differ significantly for the various delay times. The number of underestimations is significantly greater than the number of overestimations.

These results confirm the finding of Experiment I. They parallel those of the authors' earlier experiments in the sagittal plane and agree in part with the findings of others (3, 4, 5, 6) in that mean errors, in every case, are underestimations of the vertical; but they disagree in that there is no appreciable change or shift in either the mean error or the frequency

of underestimations with changing delay time. The consistent underestimation appears to be a function of the magnitude of initial tilt and the slow velocity of righting.

TABLE IV

MEAN ERROR, S.D. OF ERROR, AND FREQUENCY OF VARIOUS RESPONSES
(+ = overestimation, − = underestimation, 0 = correct estimate)

Delay time (in seconds)	Direction of error (frequencies)			Mean error	S.D. of error
	+	0	−		
0	4	—	6	−0.38	2.52
5	2	1	7	−1.10	1.90
10	2	1	7	−0.73	1.71
15	3	—	7	−0.68	2.23
30	2	—	8	−1.98	3.06
45	2	—	8	−1.83	2.80
60	2	—	8	−1.83	1.98
75	2	2	6	−1.23	2.65
90	1	—	9	−2.18	1.40
105	3	1	6	−0.93	2.35

EXPERIMENT III

Subjects were tilted 15.0° to the right at a speed of 12.0° per minute and were held in that position for 15 seconds before being returned toward vertical at 3.5, 6.5, 14.0, 17.5, 19.0, 22.0, 25.0, 28.0, 31.0, and 34.0 degrees per minute.

Results and Discussion

Table V shows the results of an analysis of variance of the data from Experiment III. This analysis reveals highly significant differences in degree of error attributable to subjects and significant differences attributable to velocity of return.

TABLE V

ANALYSIS OF VARIANCE OF ERROR BY SUBJECTS AND VELOCITIES

Source	d.f.	Sums of squares	Mean squares	F-ratio
Subjects	9	422.0025	46.8892	16.60**
Velocity of return	9	65.7775	7.3086	2.59*
Error	81	228.7350	2.8239	
Total	99	716.5150		

**Significant at the 1% level. *Significant at the 5% level.

Table VI gives the mean errors and S.D. of error for all subjects for each of the ten velocities. It can be seen that mean error decreases as velocity of return increases.

These data confirm the finding of Experiment I that increased velocity of return decreases the error of underestimation. They are similar to those previously reported for sagittal tilt and are in agreement with Fleishman's (2) findings.

TABLE VI

MEAN ERROR AND VELOCITY OF RETURN

Velocity of return	Mean error	S.D. of error
3.5°/min.	−3.40	2.60
6.5°/min.	−2.50	2.35
14.0°/min.	−1.78	2.17
17.5°/min.	−1.65	2.34
19.0°/min.	−1.28	2.11
22.0°/min.	−2.20	2.34
25.0°/min.	−1.50	2.50
28.0°/min.	−1.10	3.30
31.0°/min.	−1.20	2.54
34.0°/min.	−0.30	3.00

EXPERIMENT IV

Sex differences as reported by Witkin (7) suggested that it would be worth while to repeat Experiment I using female subjects.

Results and Discussion

An analysis of variance revealed that the effects of the four variables were the same for female as for male subjects. However, mean errors made by female subjects were significantly smaller than those made by male subjects.

SUMMARY AND CONCLUSIONS

Experiments with a tilting chair apparatus evaluated the effects of four variables upon the error of perception of the postural upright in the absence of visual cues. The variables studied in the lateral plane were: the degree of initial tilt, the time held at initial tilt, the direction of tilt, and the angular velocity of return. Female subjects were tested in one experiment. From the results of these experiments it may be concluded that:

1. The non-visual perception of the postural upright when moving toward the vertical from positions of right or left tilt in the lateral plane,

is influenced by the magnitude of initial tilt, the speed of return, and the direction of initial tilt. Errors in perception of the vertical increase as the magnitude of initial tilt increases, as the velocity of return decreases, and when the subject is returning from right rather than left tilt;

2. Subjects differ significantly in their ability to estimate the vertical. Females make significantly smaller errors than males;

3. The length of time in a tilted position preceding righting has no significant effect upon the magnitude and direction of error;

4. In all essential points these conclusions agree with those previously reported by the authors for tilt in the sagittal plane (1).

REFERENCES

1. Clegg, W. C., & Dunfield, N. M. Non-visual perception of postural vertical: I. Sagittal plane. *Canad. J. Psychol.*, 1954, 8, 1–9.
2. Fleishman, E. A. The perception of body position. Effects of speed, magnitude, and direction of displacement on accuracy of adjustment to an upright position. *Human Resources Res. Center Bull.*, 1953, 53–1. San Antonio, Texas: Lackland Air Force Base.
3. Mann, C. W., & Passey, G. E. The perception of the vertical: V, Adaptation effects. *Joint Project Rep.*, No. 9. Tulane Univer. of Louisiana under contract N7 onr–434–T.O.I. ONR project designation No. NR 140–455. U.S. Naval Sch. of Aviation Med. and Res. and Bur. of Med. and Surg. Project No. NM 001.037.
4. Mann, C. W., & Passey, G. E. The perception of the vertical; adjustment to the postural vertical as a function of postural tilt and duration of exposure. *J. exp. Psychol.*, 1951, 41, 108–113.
5. Mann, C. W., Passey, G. E., & Ambler, R. K. The perception of the vertical: VII, Effect of varying intervals of delay in a tilted position upon the perception of the postural vertical. *Joint Project Rep.*, No. 12. Tulane Univer. of Louisiana under contract N7 onr–434–T.O.I. ONR project designation No. NR 140–455. U.S. Naval Sch. of Aviation Med. and Res. and Bur. of Med. and Surg. Project No. NM 001.063.01.12 (formerly NM 001.037).
6. Passey, G. E., & Guedry, F. E. The perception of the vertical: II. Adaptation effects in four planes. *J. exp. Psychol.*, 1949, 39, 700–707.
7. Symposium. *Psychophysiological factors in spatial orientation.* Pensacola, Florida, October 1950, NAVEXOS P–966.
8. Teuber, H. L., & Mishkin, M. Judgment of visual vertical after frontal lobe injury. Personal communication, reference DRML 134–33.

CIVILIAN TO SOLDIER: THREE SOCIOLOGICAL STUDIES OF INFANTRY RECRUIT TRAINING[1]

DAVID N. SOLOMON

Defence Research Medical Laboratories, Toronto

DURING the summers of 1950, 1951, and 1952, the Defence Scientific Service sent research teams into the field to study some of the phenomena of human behaviour associated with the introduction of infantry recruits into the Canadian Army.[2] The purpose of these studies was to increase our understanding of these phenomena so that we might offer military authorities findings and recommendations that might be useful in formulating and implementing certain aspects of military training policy. Each study was reported at some length; the present paper summarizes only a few of the findings.

METHODS AND DATA

In each study observers, wearing uniform and identified by Defence Scientific Service shoulder flashes, shared quarters and messing with recruits and their instructors, participated to some extent in training activities, and recorded their observations daily. During the 1952 study, each of 35 recruits was also interviewed four times during the first twelve weeks of training, their officers and NCO's somewhat less frequently. The illustrative material in the following sections is drawn from observers' records and interview protocols. All names are fictitious.

1950 Study

This first exploratory study, in addition to making various recommendations, drew attention to an important aspect of the problem of incorporating recruits. This was stated as follows:

> A crucial phase of fitting new personnel into any organization is precisely the matter of developing in them the sort of self which corresponds to the type of organization. . . . The critical change involved in becoming a soldier is the acceptance of an altered view of oneself. Whereas one had previously viewed himself as a civilian, with appropriate habits, tastes, reactions to authority, and so on, he now needs to view himself as a participant in army life, with its distinctive requirements. This

[1]D. R. B. Project No. D 77–94–65–07. This paper was read at the Annual Meeting of the Canadian Psychological Association, Kingston, Ont., June, 1953.

[2]Dr. Oswald Hall, Chairman, Dept. of Sociology, McGill University, has been consultant to this project. Other participants were: Dr. W. E. Westley, Messrs. J. E. Brazeau, F. E. Jones, B. A. MacFarlane, T. F. S. McFeat, H. W. F. McKay, T. E. Rashleigh, and H. Rodman; and Miss A. J. Wipper.

amounts to a transformation, which involves shedding many deeply ingrained habits and tastes, and incorporating and internalizing many distinctive new habits and attitudes. The change extends deeply into the sentiments and feelings which the soldier experiences and expresses. . . . The core of the matter is thus bringing the person to view himself as a soldier rather than as a civilian. When he gets to that point, the proper behaviour for a soldier becomes spontaneous and largely automatic.

The six statements which follow are drawn from over 100 in the report. Except for the remarks of the lieutenant-colonel, all occurred during conversations in which the observers were included. They illustrate the type of material which suggested that there was a difference between men who possessed something like the appropriate self concept and those who did not.

Private soldier: When I first came into the Army I surrendered completely to it. I can't understand other men who come into the Army and don't have the feeling of wanting to take in everything the Army has to offer.

Another soldier: I have a job already waiting for me with the Bell Telephone. My girl friend's father has it for me . . . the PSO said he'd recommend me for discharge.

A sergeant: I went down to headquarters in Ottawa. But I stayed only five days. After being there three days I found myself walking down the hallways with my hands in my pockets. I told myself that this wasn't my true self. So the rest of the time I was there I used to carry a swagger stick in my hand and I marched down the hallways. Of course, all the people who saw me were quite amazed. When I went before the colonel down there to ask to be sent back here to the regiment, I told him when I wanted to join the Civil Service I wouldn't do it through the Army.

A corporal: I'm leaving at the end of my three years . . . I'm going back home to London to take over my father's farm. I always liked farming. It's like having your own business. In the Army it's different, you don't own your own business.

A lieutenant-colonel (to a meeting of battalion officers): Shoes will be simple, no brogues . . . There will be no gaudy ties with mufti. All officers will have blazers. When the General comes, all officers will have their wives call and leave their cards. If single, the men will do the calling. These amenities are a part of the life of the soldier. I tell you this for your own good. It is a part of a career. Your wife can make or break you. If she hates army protocol, just tell her, "Honey, I'll be a lieutenant the rest of my life." Single men please listen.

A young lieutenant: Why must it [the regimental mess dinner] be every week? Same old people, same old topics of conversation, same old games. Why not get excused and go off to dinner somewhere else?

Another young officer: Utterly boring. Too hot, too many mosquitoes, and too often.

These three sets of quotations illustrate the contrast which impressed the research team. There appeared also to be a middle group, not represented here, who intended neither to make a career nor to obtain a discharge, but to lessen their discontent by obtaining a transfer to another company or corps. Although the picture was not clear-cut, there was a

general contrast between the devotion to military life of senior NCO's and senior officers, and the tenacious attachment of some young soldiers—both officers and other ranks—to a civilian way of life. The former seemed to have undergone, in the course of time, a transformation which led them to see themselves and their relation to the service in a different light.

1951 Study

Central to the notion of self-conception put forward above is the belief that it is formed in personal interaction with others.[3] The 1951 study was accordingly focussed on the immediate face-to-face contacts of recruits. By chance three groups were discovered, each consisting of about ten men with a corporal, which differed in ways that seemed to have particular significance both for those responsible for training and for the research team studying the "atmosphere" of small training groups.

Group 1. One observer lived with this group for the first four weeks of their army life. The corporal spent much time with them, both during training and after duty. Through constant association all became well acquainted, and each knew something of the previous history of the others, their families, girl friends, occupations, and the like. Although the corporal firmly demanded high standards in training, he also saw conscientiously to welfare—helped the men get uniforms and kit, find their way to sick parade, or to an officer when family trouble arose. These men thought of themselves as a group. They thought of the corporal as *our* corporal, and he thought of them as *my* men. From the army point of view this was an ideal situation, where interest and progress in training appeared much better than average.

Group 2. The second group was observed for a somewhat shorter time. These men, and our observer, also lived together with a corporal. The corporal, however, took little interest in the men, and his contacts were mainly disciplinary. He had little responsibility for training, since they moved to a different instructor for each subject. The men seemed to have less interest in each other, and there were no strong reciprocal bonds between them and the corporal.

A Test of "Morale." Quite fortuitously these two groups were tested by a storm, which threatened to flatten the tents and drench the men and their possessions. Both corporals were away at the time; both groups had experienced or heard of similar storms since arriving in camp. In the first

[3]See, in this connection, Cooley, C. H., *Human Nature and the Social Order* (New York: Scribner's, 1902), pp. 183–184; Mead, G. H., *Mind, Self and Society* (Chicago: Univer. of Chicago Press, 1934), pp. 153–154.

group, one or two men quickly assumed leadership and mobilized the group, so that by the time the corporal returned they had done almost everything necessary to protect the tent and keep themselves dry. The second group did nothing, even though they observed the approach of the storm. When the wind and rain struck they still did nothing to protect the tent or their clothing and equipment, until our observer, unable to remain neutral any longer, encouraged them to take action, which they did with little enthusiasm. Throughout the storm they remained largely passive, and later became somewhat resentful. The men of the first group, in contrast, looked on the storm as a challenge, and felt they had exhibited their toughness in coming through it with kit and spirits undampened.

The comparison of these two groups highlighted the importance of the role of the corporal, who represented the Army to the young recruits.

Group 3. Another aspect of the recruit training situation emerges from a third observer's description of his group, who also had lived and trained together for about three weeks. After duty they often discussed training, demonstrating the rifle to each other, showing each other how to fix berets and belts, and working together making beds and preparing kit. They got to know each other well, and regularly went to the movies or the canteen together. Some balked at training, "beefed," and talked about going "AWOL"; the others, however, vigorously supported the Army, and shouted down the gripers. They tended to put pressure on deviants to smarten up drill, uniforms, and kits, by jeering and cursing at them and complaining that they were holding the whole group back.

After about three weeks these men were separated and distributed through nine different tents, although they continued to train as a group. In their new tents they exchanged brief life histories, began to question the older men about the *minutiae* of training, and occasionally played cards. But the new tent-groups lacked the solidarity of the old. No one bothered to shout down gripers, tent-mates seldom helped each other with problems of training and housekeeping, and after duty they spent little time together. However, the members of the old tent-group who still trained together met regularly. They frequently gathered in one or another of the tents, and went on pass together. They played cards, talked about women, and griped; but they no longer discussed training or tried to assist each other.

In short, when these men lived and trained together they formed a close friendship group which coincided with army structure. The solidarity of this group facilitated the training programme in many ways. When the tent-group was broken up the friendship continued, but at a lower level of intensity, and it no longer coincided with army structure.

The joint activities were leisure-time ones, and made little contribution to the training programme.

This suggested that the nature of informal interaction among recruits was also important for the outcome of training.

1952 Study

The 1952 study, accordingly, sought more detailed information on the interaction of recruits with (*a*) authority figures, such as corporals, and (*b*) other recruits. This paper will omit informal, peer-group interaction, and confine itself to interactions in certain formal situations.[4]

In the training programme there appear to be three types of formal situation, namely, the parade square, lectures, and field training exercises. Although all are formal in the sense that activity is consciously directed and coordinated, they differ in the degree of insistence that all the actors play the roles indicated for them by the rules of the larger formal organization of which the training group forms part.

1. *Parade Square.* This is the first formal situation to which recruits are introduced. It is mainly disciplinary in character. Officers are present infrequently, usually for an inspection or in some similar role. The sergeant is the protagonist of the parade square situation, seeing himself, as the following interview excerpt indicates, as the custodian of discipline.

Sergeant: The main thing I'm interested in teaching them is obedience. Now in the first three weeks that is what I'm trying to do. I want them to get the habit of listening to me when I speak and to do what they are told. When I begin to instruct or come back after a break, as soon as I begin to say something I want absolute silence and not a lot of talking between them or looking around and all that sort of thing. I want them to pay attention and they have got to get the habit. They have got to know that when a parade is called for 0800 that they have to be there at 0800 waiting for me. I don't want to have to go around the huts telling them they ought to be out on parade. Now the only place they can learn that is on the parade square.

The following diary accounts of parade square occurrences illustrate the kind of discipline which the sergeant prescribes and the corporals enforce.

Lance-Corporal: All right now, what's so funny? Stop that laughing or I'll parade you under escort.
Recruit: I wasn't laughing.

[4]Formal situations are those in which activity is more or less consciously coordinated and directed, as distinguished from those in which interaction is spontaneous, free-flowing, and relatively undirected. See also Barnard, Chester I., *The Functions of the Executive* (Cambridge: Harvard Univer. Press, 1951), p. 73.

Lance-Corporal: Don't talk back.
Recruit: I wasn't laughing . . .
Lance-Corporal (shouting): Shut up!

Riley, attempting to fall in in the middle of a file, pushed another man, and a brief scuffle occurred. Lance-Corporal Forrest came over and asked what had happened, and Riley started to explain.

L/Cpl. Forrest (interrupting): Never mind.
Riley: ——you!
Forrest: You're on charge.
Riley: I don't care if I'm on ten —— charges.
Forrest: You're under close arrest.
Riley: I don't give a ——.
Forrest: Two men fall out!

Riley was charged and later admonished. Threats of charges were, however, more frequent than actual charges, and for the most part recruits responded to discipline compliantly.

2. *Lectures.* The lecture situation, to which recruits are also introduced during their first few weeks, contrasts markedly with the parade square. The emphasis is on instruction rather than discipline; although the corporals are still superordinate, the situation is more permissive. There is a certain amount of joking and the recruits have occasional opportunity to initiate interaction, usually by asking a question, not always entirely relevant.

Riley: Corporal, how many moving parts on a bren?
Cpl. Landon: Why?
Riley: Well, I had 'em all moving a couple of days ago.
Weil: You crazy ——.
Riley: Who are you referring to; me or the corporal?
Cpl. Landon: Aw cut it out, Riley.

The questions and fooling continued for some time, so Landon said, "OK, quieten down. There's some guys who want to listen and learn something. So we'll keep quiet, OK? OK!" He described the gun, dismantled it, and carried on with the lecture and demonstrations. Landon caught Weinberg sleeping, and made him stand up. Later he caught him a second time, and said, "Are you sleepy, or is it hot in here, Weinberg? Just let me know, that's all." Then he propped the door open. Later Pace was sleeping, and Landon said: "For —— sake, wake up. Right in front of the class; don't fall asleep." Later, Pace knew the name of a part of the gun which the corporal himself had forgotten. Someone said, "You can go to sleep now, Pace," and everyone laughed. Again Corporal Landon noticed him dozing. He quietened the class and waited to catch him. Pace awoke and everyone laughed. Towards the end of the lecture an officer passed the open door. Corporal Landon saw him, and motioned the class to quieten down. They did, and started to ask questions. Landon kept looking towards the door and motioning towards the class to keep quiet.

The contrast between this and the parade square is obvious. In the latter all interaction is initiated by sergeant or corporal, and recruits are

obliged to respond exactly as indicated. In the lecture situation, however, there is some relaxation of the formal pattern. Recruits may be encouraged to ask questions, and even permitted to initiate joking and banter. Whereas the context of the parade square is essentially disciplinary, that of the lecture is more like the atmosphere of a classroom—partly disciplinary, but mainly oriented towards instruction, and permitting some degree of friendly, spontaneous interaction between those of near-equal rank.

3. *Field Exercises.* Field exercises presented a further contrast. Their attempted simulation of combat appealed to recruits. However, the main contrast was that in field exercises the recruits, who had previously been responding, along rather clearly delimited lines, to actions initiated by others, were now to a considerable extent participants in a team effort. For the first time the work of officer, sergeant, and corporal was visible to the recruits, and they could see the working of the chain of command which linked their own activity to the battalion goals. They saw also that the functions of their superiors were not merely disciplinary or punitive, but included providing for the welfare of recruits—rations, water, seeing that they were bedded down as comfortably as possible, and so on. To some extent all suffered the same hardships, and all had an interest in solving both tactical problems and those presented by outdoor living.

The following diary excerpts indicate the team character of the approach to problems of tactics and housekeeping.

Upon arrival the recruits were ordered to dig in. The lance-corporal pointed to a spot where he wanted a slit trench, just at the edge of a bush.

Recruit: We won't be able to see from there very good, corporal; I think we should dig it out from the bush some.

Lance-Corporal: You've got to be hidden from the air.

Recruit: That's all right. I'll camouflage the hole so that no one can see it.

Lance-Corporal: All right, but I'll come back and check it, so it better be good.

Cpl. Donalds said we should all go for wood to keep the fire going all night. Adrien was the first to go, and he suggested, after he and Cpl. Donalds got the fire going, that we set up a crosspole over the fire and use it to dry our socks and puttees. Pace suggested that we use rifles, with bayonets fixed, as end poles.

These descriptions illustrate the kind of experience recruits had with three levels of formal organization during the first ten to twelve weeks of training under the policies in force during the summer of 1952. Each situation had a different emphasis: discipline on the parade square, a degree of permissiveness in the lecture room, and teamwork during field exercises. In each case instructors acted in markedly different fashion toward recruits. Moreover, while initially recruits spent much of their

time on the parade square, or in similar situations, as time went on lectures assumed greater importance, and finally field exercises of several days' duration became frequent. Thus recruits learned how to respond first in one situation and then in another. The recruit's status may be thought of as changing from that of a person treated very strictly, who cannot be relied upon for independent action, to that of one deserving a certain amount of lenience, and finally of one who is called on to take a considerable amount of independent action, and is treated as a responsible individual.

Discussion

This paper summarizes a few aspects of three extensive studies prepared for military readers. The phenomena reported are not integrated by any tightly knit conceptual scheme. The reader must assess the continuity in the insights and leads followed by the researchers, as they proceeded from social psychological phenomena, such as the self concept, to the structural aspects of the training situation.

Beginning with the important contrasts between those who had come to conceive of themselves as soldiers and those who had not, we were led to study the small-group experience of infantry recruits, on the assumption that the self-concept is formed in face-to-face interaction. There the significance of the NCO's role was manifest, and led us to study three situations characterized by differences in the way in which recruits were treated by their superiors. In these, recruits appeared to be learning progressively how to redefine themselves and their roles.

As we have noted, most recruits conform well to what is expected of them. However, it is difficult to do otherwise in the Army, and we have only fragmentary evidence of the extent to which such conformity represents a changed conception of the self, an internalization of the Army's demands. Some recruits, when asked how "the recruits" got along on an exercise, replied that no recruits had been on the scheme. It turned out that a new batch of recruits had arrived recently and had not participated. Our men, who a fortnight earlier would have referred the question to themselves, had shifted themselves to a new category; some termed themselves "senior recruits." There are other suggestive passages in our data; further analysis and research may reveal how far such felt changes in status represent internalization of the values of the soldierly way of life.

ON LEARNING AND HUMAN ABILITY[1]

GEORGE A. FERGUSON
McGill University

MY PURPOSE is to present a generalized theory which draws together within a single conceptual framework the study of human learning and the study of human ability. Those concerned with the description and classification of man's abilities have usually adopted an individual difference approach. They have paid scant attention to problems of learning. The experimentalists, engrossed in the study of learning, have for various theoretical and practical reasons shown little interest in individual differences. They seem unaware that they too are students of man's abilities. This divergence between two fields of psychological endeavour has led to a constriction of thought and an experimental fastidiousness inimical to a bold attack on the problem of understanding human behaviour.

At present no systematic theory, capable of generating fruitful hypotheses about behaviour, lies behind the study of human ability. Current approaches are largely empirical. Psychological test theory and factor theory, evolved as they were for the study of human ability, are largely technologies which do not presume to answer psychological questions *per se*, although they may aid in the answering of such questions once raised. The pioneers of factor analysis, Spearman (**16**), Thomson (**17**), and others, proposed theories of brain functioning, and resorted to factor analysis as a means of testing deductions from those theories. Few attempts are made today to correlate the descriptive parameters of behaviour identified by factorial methods with any structural or dynamic properties of a brain model. Many factorists, although accepting this as a legitimate problem, regard its exploration as beyond their province and possibly premature. Earlier factorists thought otherwise. Many physiological psychologists on the other hand still concern themselves with rather vague and global concepts of intelligence, and seem unaware that these concepts are regarded as obsolete by many students of human ability. The concept of intelligence, however it be framed, is no longer a useful scientific concept except as subsuming some defined set of clearly distinguishable abilities.

If the study of human ability lacks theoretical buttressing, the study of human learning most certainly does not. The physical bases of memory have been the subject of much speculation at least since the time of René Descartes. One of the postulates of a recent theory proposed by

[1]D. R. B. Project No. D 77–94–01–01.

Hebb (10) is identical in form with the postulate of Descartes. The extent and persistence of this class of theorizing are apparent in a recent extensive review by Bronislaw Gomulicki (8). One recent line of speculation in this area has been advanced by the cyberneticists and by those concerned with the development of electronic calculators and related devices. Analogies have been drawn between the principles which govern the functioning of machines that "learn" and "remember" and the principles which govern the functioning of the human brain.

In the field of human learning there is clearly a plenitude of diverse theoretical constructs, together with an extensive accumulation of experimental data. In the field of human ability, although we suffer from a paucity of systematic theoretical constructs, we have an extensive technology, much data based on the study of individual differences, and formal ways of thinking about problems which are, I believe, foreign to many who work in the learning field. It follows that, if we can logically incorporate the two fields in a single conceptual framework, some mutual enrichment may occur.

Ability and Overlearning

"Ability" is defined operationally by the performance of an individual in a specified situation. Thus L. L. Thurstone (18) states that "an ability is a trait which is defined by what an individual can do." It follows, as Thurstone points out, that "there are as many abilities as there are enumerable things that individuals can do." Factor analysis, a classificatory technique, undertakes a parsimonious description of the multitudinous array of abilities in terms of a relatively small number of categories. It is clear that the term "ability," in addition to its operational meaning, may be assigned a formal postulational meaning within a framework of theory.

"Learning," as conventionally used in experimental psychology, refers generally to changes, with repetition, in "ability" to perform a specified task, the changes being regarded as functionally dependent on, or in part assignable to, repetition. Other assignable causes of change, such as fatigue, sensory adaptation, artifacts of measurement, and the like, are presumed to be controlled. This commonly accepted statement of what is meant by learning is inadequate and leads to logical pitfalls. It will, however, serve my immediate purpose. Note that I have introduced the term "ability" into the definition of learning. It is seldom used in connection with learning experiments, although indices of performance used in such experiments are clearly measures of the ability of the subject at various stages in the learning process. Conventional learning curves are simply descriptions of changes in ability with repetition.

If we regard the term "ability" as defined by an individual's performance, we may identify two broad classes of ability: a class which is more or less invariant with respect to repetition or its cessation, and a class which is not. Thus some of the things which individuals can do appear to have a fairly high degree of permanence, showing little change, either in the presence or in the absence of repetition. Other things individuals can do may exhibit gross improvement with repetition, or gross impairment following a period in which no repetition occurs.

The typical learning curve shows that in most learning situations a level of performance is attained with repetition beyond which no further improvement is observed. Moreover, in certain learning situations, particularly those demanding a high degree of overlearning, the subject when tested may exhibit no impairment in performance even after lengthy periods of time without repetition. It seems that in most learning situations the ability of the subject reaches a crude limit[2] beyond which no systematic improvement is likely to occur with repetition. Of course in some cases rapid improvement may occur following a "plateau" in learning.

While the term "ability" may be used generally to refer to performance on any type of task, whether or not it varies with repetition, many psychologists conventionally use the term, in the case of adults, to refer to performance which does not vary much over lengthy periods of time. Likewise it is assumed that the abilities of children are reasonably stable over short time intervals, although they may show systematic improvement with age. For example, in Thurstone's classification of mental abilities (**19**), reasoning ability, number ability, perceptual ability, spatial ability, and the like, are presumed, in the adult subject, to be reasonably stable attributes of behaviour over lengthy periods and, in the child, to have considerable stability at any given age level. It is not presumed that an individual's reasoning ability will be markedly changed by solving large numbers of verbal analogies or number series items, or that his spatial ability will show pronounced improvement with practice on paper form-board items or other spatial tasks. Although some improvement with practice will occur in many subjects, we do not expect such improvement to be gross.

These observations bring us to the first hypothesis of this discussion, namely, that in the adult subject in our culture, those more or less stable attributes of behaviour which we ordinarily speak of as abilities, and

[2]The term "limit" is used here in a very loose and imprecise sense to refer to a level beyond which no very gross and systematic change seems likely to occur. The term is not used in the rigorous sense in which it is understood in elementary mathematics or in the stochastic sense of probability theory.

which are defined in terms of performance on psychological tests, refer to performance at a crude limit of learning. This is regarded as applying to all attributes in the Thurstone classification and to whatever is subsumed under the term "intelligence." The hypothesis implies that these abilities are overlearned acquisitions, and that the stability which characterizes them is the result of overlearning. It assigns to learning a central role in the study of human ability, and opens the way for the study of ability and learning within the same conceptual framework. It is proposed that any theory which implies that individual differences in ability are individual differences at some crude limit of performance reached by overlearning be referred to as a *limits of learning* theory of human ability.

The role of learning in human ability is well illustrated by number ability. Number ability is defined by the performance of individuals on certain simple arithmetical tasks involving addition, subtraction, multiplication, and division. Tests of number ability are usually highly speeded. In our culture the majority of educated adults are fairly facile at tasks entailing ordinary arithmetical operations. Arithmetical facility has for many individuals been so reinforced by innumerable repetitions over prolonged periods of time that a crude limit of learning has been attained. Although some might show improvement with systematic practice, this would probably not be great. It seems plausible, therefore, that in many adults individual differences in number ability refer to individual differences at some crude limit of performance.

The role of learning in perception is a controversial subject. Much evidence supporting the view that perception involves a prolonged period of learning has been marshalled by Hebb in *The Organization of Behavior* (**10**). This evidence strongly suggests that various perceptual abilities represent performance at the limits of learning in perception. In the normal child the limits of learning in many perceptual tasks may be reached at a fairly early age.

The role of learning in reasoning—that is, in the type of ability required for the solution of number series items, verbal analogies, and so on—is far from obvious. But numerous arguments can be advanced to support the view that reasoning ability involves a prolonged period of learning.

The view that "ability" has reference to performance at some crude limit of learning is not new. It is implicit in the theories of Hebb (**10**) although not formulated by him in the above manner. In much of the work carried out in the animal laboratories at McGill the "intelligence" of the rat or dog is defined in terms of performance on a maze test following a lengthy learning period. The animal performs in the test situation until his performance reaches a crude stability, or until a limit of per-

formance is very roughly approximated, and this limit serves to define his "intelligence."

Gross individual differences in ability do exist. These differences are a complex result of the interaction of the biological propensities of the organism and the learning which occurs at particular stages of life. This topic has been discussed by Hebb (**10**). His distinction between early and late learning is relevant here. It appears that the state of an organism at any given time and its ability to respond to any immediate situation is a complex function, not only of its biological propensities and previous learning, but also of the stage in life at which learning of various types has occurred. This implies that the stage of development at which learning of a particular class has occurred is one factor in determining the limit of overlearning at the adult stage. Thus, as Hebb observes, early learning or its lack may have a permanent and generalized effect in the adult.

Many of the abilities which psychologists have studied increase with age. Intelligence as defined by such tests as the Stanford-Binet increases until about the age of 17, when a limit of performance is reached. In our culture children are exposed to an environment that demands rapid learning of many things. They proceed as rapidly through the school system as their abilities at any stage will allow. It is probable that many children at any particular age are functioning fairly close to the limit of their potentiality with regard to certain classes of activity. It follows that some of the abilities measured by psychologists are, for many children, indices of performance at a crude limit of learning for the age in question. However, if a child's environment is restricted with respect to certain activities he may function well below the limit of his potentiality in those activities at varying ages, and a permanent impairment at the adult stage may result. Presumably children reared in different environments, which demand different types of learning at different ages, develop different patterns of ability.

Transfer

"Transfer" is frequently used in a general sense to refer to the effects of changes, resulting from repetition, in ability to perform a specified task, on the ability to perform either the same task under altered conditions or a different task. This is the meaning usually assigned to the term "transfer" for laboratory experimentation. An implicit condition is that the prior task is in some respect different from the subsequent task. When the two tasks are presumed on the basis of superficial inspection to be similar, the term "learning," and not "transfer," is used to refer to the changes in ability that occur. It seems therefore, that transfer is the general phenomenon and "learning" is a particular *formal* case which may never

occur either in laboratory experimentation or in real life situations. The notion of learning implies the identity of a sequence of learning situations.[3] The fact that learning is a particular formal case of the general phenomenon of transfer has been recognized by Cook (6) who writes:

> There is no separate problem of transfer of training. Or conversely, all learning (unless there exists a limiting case in which successive trials are identical on all counts) involves the problem posited in the transfer of training experiments: What identities and differences in successive trials affect what sort of learning?

With the possible exception of some learning which occurs very early in life, all learning occurs within the context of experience. We bring to bear on the learning of any task a mass of prior experience which may either facilitate or inhibit the learning of that task. On this point McGeoch (12) writes:

> After small amounts of learning early in the life of the individual every instance of learning is a function of the already learned organization of the subject; that is all learning is influenced by transfer. . . . The learning of complex, abstract, meaningful materials and the solution of problems by means of ideas (reasoning) are to a great extent a function of transfer. Where the subject "sees into" the fundamental relations of a problem or has insight, transfer seems to be a major contributing condition. It is, likewise, a basic factor in originality, the original and creative person having, among other things, unusual sensitivity to the applicability of the already known to new problem situations. Perceiving, at whatever complex level, is probably never free of its influence, and there is no complex psychological event which is not a function of it.

---

[3]The inability of students of learning to deal appropriately with the problem of the identity of and difference between tasks has led to logical difficulties in our concepts of learning and transfer. To my mind these concepts require some revision. To say that an individual is repeating the same task, or that one task is different from another, demands a precise statement of what is meant by "same" and "different." The referents of these terms in current thinking on learning are largely phenomenological; that is, they have to do with our immediate experience upon inspection of the tasks in question. Two ways out of this difficulty suggest themselves. First, the terms "same" and "different" are always with respect to some property or properties. In view of this, it may be possible to define operationally properties with respect to which tasks may differ, and to study the relationship of such differences to differences in transfer effects. Attempts have been made to do this. Second, "same" and "different" may be defined in terms of correlation. If the correlation of the performance of a group of individuals on two tasks, or on successive trials of what experientially is the same task, is roughly unity, error being taken into account, then the tasks may be said to be the same. If the correlation departs from unity then the tasks may be said to be in some degree different. We may be prepared to go beyond a strict operational statement of this kind and speak of tasks on successive trials as involving the same or different functions. This essentially is the rationale of factor analysis. This second approach in effect defines the stimulus in terms of the responses of the subjects.

Hebb (**10**) in discussing this same point writes:

> If the learning we know and can study, in the mature animal, is heavily loaded with transfer effects, what are the properties of the original learning from which those effects come? How can it be possible even to consider making a theory of learning in general from the data of maturity alone? There must be a serious risk that what seems to be learning is really half transfer. We cannot assume that we know what learning transfers and what does not: for our knowledge of the extent of transfer is also derived from behaviour at maturity, and the transfer from infant experiences may be much greater and more generalized.

If all adult learning is heavily loaded with transfer, what is the nature of the prior learning which transfers to the learning of new tasks, and how does it affect the learning of such new tasks? Two hypotheses are put forward.

The first is that, in many adult learning situations, the most important variables exerting transfer effects on subsequent learning are the "abilities" –the prior acquisitions that have attained their limit of performance. This hypothesis has long been widely accepted and is deeply entrenched in our thinking. It is commonplace to say that "bright" children learn more quickly in school than "dull" children, where brightness and dullness are defined in terms of performance on an ability test. The validation of tests against training criteria implies that the abilities of man are significant variables in the learning process. Such a loose statement as "Intelligence is learning ability" reflects the important role frequently assigned to intelligence in learning situations. There are two reasons for emphasizing this hypothesis. First, it is formulated within a new theoretical framework which alters substantially our way of thinking about the role of human abilities in learning. The problem can now be regarded as a problem in transfer. Second, the role of human ability in human learning has always been a matter of major practical concern to the applied psychologist. Indeed, it may be the most important problem in the applied field. And yet this problem has received little attention from the theoreticians or the laboratory experimentalists. Experiments on transfer carried out under laboratory conditions are so distantly removed from learning as it occurs in real life situations that they provide few answers of the slightest usefulness in the field of applied psychology. Questions of what prior learned acquisitions, or abilities, transfer to what learning, and how, and under what conditions, remain largely unanswered.

The second hypothesis concerns the way in which overlearned acquisitions, or abilities, affect subsequent learning. It is that such abilities exert their effect differentially in any learning situation; that different abilities exert different effects at different stages of learning, and that the abilities which transfer and produce their effect at one stage of learn-

ing may be different from those which transfer and produce their effects at another stage. This means that individual differences in abilities which may be functionally related to individual differences in performance in the early stages of learning a task, may not be functionally related, or may be related in a different way, to performance in the later stages. An implication of this hypothesis is that an individual might possess the abilities to perform a given activity with a high degree of proficiency, but might lack the abilities to learn to perform the task under certain specified conditions of learning. Likewise an individual might possess the ability to improve rapidly in the early stages of learning, but might lack the abilities necessary to attain high proficiency at the stage of high habituation or overlearning. The learning of many motor activities probably belongs to this class.

This hypothesis, if experimentally confirmed, will have important educational implications. It implies, for example, that a slow learner under given learning conditions may have a capacity for ultimate performance in excess of the fast learner under the same training conditions. In the test validation field, where tests are frequently validated against training criteria, it becomes important to consider the stage of training to which the criterion relates, since tests with an acceptable degree of validity at one stage of training may have little or no validity at another stage.

A prior overlearned acquisition, an ability, may not only facilitate the learning of a new task but may also inhibit it. Thus we may consider both positive and negative transfer effects, and the simultaneous operation of such effects. Although the terms positive and negative transfer are used to refer to net effect of the operation of a variety of variables, experiments could readily be designed to separate out the positive and negative effects of different abilities on the same learning situations.

Experimental Design

The experimental investigation of these hypotheses involves an individual difference approach. There have been a number of such approaches in the learning field (**14**, **20**, **22**, **23**), but this line of attack has been relatively unpopular, owing to the practical difficulties of finding appropriate learning tasks which will provide reliable measures of performance and permit the collection of data on substantial numbers of subjects.

The type of experiment suggested by the hypotheses in this paper may be illustrated as follows. Say that we are concerned with the transfer of certain overlearned acquisitions, or abilities, to the learning of a motor

task, that a number of learning periods are allowed, and that a measure of performance, or score, is obtained for each learning period for each subject. We may select and administer to our subjects a number of tests of abilities which *a priori* considerations have led us to believe may transfer either positively or negatively to the learning of the motor task. For any group of subjects the intercorrelations between all the variables may be calculated. The relationship between performance at various stages of learning is described by the correlation between scores on the learning task. The correlations between the ability tests and scores on the learning task are measures of the extent of transfer of the abilities to the learning situation at various stages of learning. The results obtained from such an experiment can probably best be handled by factorial methods, treating the tests of ability as criterion variables.

Experiments of this design will permit observation of the differential transfer of abilities at different stages of learning. The design may also be readily extended to cover forgetting, and the differential effects of transfer through a cycle of learning and forgetting. Further, the design permits the differentiation of the simultaneous operation of positive and negative transfer effects.

Culture and Human Ability

Extensive investigations have been carried out on the effects of various cultural factors on "intelligence," as defined by the standard intelligence tests. The general conclusion is that a variety of cultural variables are related to "intelligence" as so defined. Since this is so, the view is held by many investigators that existing intelligence tests when used in selection, classification, and the like are "unfair" to certain sections of the population. This has resulted in attempts to develop types of test material which are more or less invariant with respect to certain cultural factors. Examples of this are the "Culture Free Tests" developed by R. B. Cattell (**4, 5**). A recent extensive investigation by Eells, Davis, *et al.* (**7**) attempted to isolate groups of test items which showed relatively small or negligible differences between individuals in various socio-economic groups.

Many investigators concerned with this class of problem regard "intelligence" as a basic underlying biological attribute. Methods of measuring it are, however, affected by cultural factors. How, then, can inferences be drawn regarding differences in cultural or racial groups in "intelligence" viewed as a biological attribute? There is no obvious answer to this question. One way out is to accept the hypothesis (incapable, of course, of any experimental test by existing methods) that no difference

exists between cultural and racial groups in "intelligence," and that where such differences are found they are the result of cultural factors. If this hypothesis is accepted the next step is obvious.

Tests must be constructed which are invariant with respect to certain controllable cultural factors, and so are "better" measures of the basic biological variable called "intelligence." As Turnbull (**21**) has recently pointed out in criticizing this line of argument, a "fair" or "good" test is one which shows no differences. Thus, as Turnbull remarks, the process has gone full circle. The hypothesis is accepted that no differences exist between cultural groups in biological intelligence. Tests are constructed by the careful selection of items which show no differences between groups. These are then used as evidence that there *is* no real difference.

The position described above has been widely adopted, although its logical ramifications have seldom been explicitly stated. Its basic weakness lies in a naïve concept of "intelligence," which leads to an experimental impasse avoidable only by the acceptance of an unverifiable hypothesis. Thus the existing position held by many investigators is logically untenable, and hence cannot lead to profitable research.

The theory presented in this paper enables us to regard these problems in a different light, and to formulate them in more meaningful terms. It states that the more or less stable attributes of behaviour, commonly referred to as abilities, represent performance at crude limits of learning, and that such limits are determined by the biological propensities of the individual and by cultural factors which prescribe what shall be learned and at what age. Therefore, questions raised about the role of cultural factors in human ability are essentially questions about the relationship between learning and human ability.

The obvious inference from this line of argument is that individuals reared in different cultures will develop different patterns of ability. It is substantiated by a mass of anthropological evidence. It must be so. It cannot be argued that, if no differences between cultural groups on a particular test are found, we are measuring a biological capacity in which no differences exist between groups. Nor, conversely, can it be argued that where differences are found they are the result of biological differences. Such arguments are ruled out of court in the above theoretical position. The initial problem becomes one of describing the patterns of ability which are characteristic of individuals reared in different cultural environments. The initial problem is not one of demonstrating that intercultural differences exist with respect to a particular ability, or that they do not, or of drawing inferences from such findings one way or another.

I regard the inferences drawn from many of the extensive investigations

on racial differences in "intelligence" as essentially invalid. Many of these studies sought differences in "intelligence," biologically regarded, between racial groups; race being a biological concept. Hypotheses pertaining to this problem are unverifiable in terms of the present theory. Racial groups may exhibit different patterns of ability, as defined by performance on particular tests, but to argue from these data for or against the existence of biological differences is not meaningful.

In Canada during the war the Armed Services attempted to develop parallel tests for French- and English-speaking personnel. The hypothesis underlying the development of such tests was that no differences in the pattern of abilities existed (or should exist) between these groups, an hypothesis which was politically expedient regardless of its scientific validity. This resulted in statistical manipulation to ensure that the French and English forms, when applied to samples of the respective populations, gave the same means and variances. A practicable approach to this problem, which avoids such difficulties, is to develop tests, possibly quite different, for French and for English, and to validate them separately in French and English situations. If similar non-language tests are developed for French and for English, and these then show pronounced differences in the patterns of ability between French and English, the problem is not one of obscuring these differences by statistical manipulation, but of ascertaining the best use to which these differences can be put in the selection and classification of personnel.

As for isolated, underprivileged, and restricted,[4] cultural communities, the initial problem is again one of developing tests for describing adequately the abilities of the members of such communities. These tests should conform to the usual criteria of reliability, discriminatory capacity, and the like. In the Newfoundland outports, for example, many individuals display excellent skill in boat building, navigation, and fishing. On tests of the abstract thinking type developed for use in urbanized cultures, the members of such communities make low scores. The tests discriminate poorly. The abilities that develop among the members of such communities, and upon which their survival depends, are probably quite different from those that our urbanized culture fosters. The first step is to discover adequate ways and means for describing these abilities. Once this is done, an attack can be made on other problems.

[4]Implicit in the use of such terms as "restricted," "underprivileged," and the like to refer to cultural groups is the evaluation that because many cultures are different from our own they are in some vague sense "not as good." Even some of our better scientific thinkers seem incapable of observing a difference between cultures without implying a value judgment. Such terms have meaning only in relation to some particular criterion variable. If that variable is the availability of medical services or

Age and Ability

There is a substantial body of experimental data on the relationship between test performance and age in different cultural groups, but it has not been adequately interpreted or carefully assessed in relation to a theory of mental ability. The theory proposed in this paper permits an interpretation of these data and leads to certain hypotheses which, if substantiated, may be of some practical consequence in the development of tests for different cultural groups.

Take the studies by Gordon (9)[5] on canal-boat and gypsy children in England. The canal-boat children received a very limited education. The average school attendance was estimated at only 5 per cent of that in the ordinary elementary schools. Each family led a relatively isolated existence and had little contact with other canal-boat families. In a sample of 76 children the average IQ (Stanford-Binet) was 69.6. Notable was the sharp decline in IQ with age. The correlation between IQ and age was —.755. The four- to six-year group had an average IQ of 90, whereas the oldest group averaged 60. In children of the same family a consistent drop from the youngest to the oldest was observed. The mental ages of children within a single family were similar, although chronological ages differed.

In the case of gypsy children the mean IQ in a sample of 82 was 74.5, and the correlation between age and IQ was —.430. The school attendance of the gypsy children was about 35 per cent of possible school days. Although IQ was negatively correlated with age, it was positively correlated with school attendance. The increment of mental with chronological age was far below that which generally obtains.

Similar results have been found in studies of mountain children made by Hirsch (11), Asher (3), Sherman and Key (15), and others. Studies on Eskimo children carried out by Anderson and Eells (2) report the same findings. The Mean IQ on the Stanford-Binet at the 8-year age level was 99.6 and at the 18-year age level 66.8. On the Goodenough scale for "drawing a man" the corresponding average IQ's were 100.0 and 87.2. Similar observations were made by Porteous (13) in applying

---

the number of refrigerators per 1,000 population, then these terms may be assigned a precise meaning in relation to such variables. If, however, the criterion variable is a phenomenological one, such as "happiness," then the terms are probably meaningless, because propositions relating to the relative "happiness" of peoples in different cultures are unverifiable. Some observers infer that the Eskimos in Baffin Island are a very "happy" people, whereas others seem prepared to contend that the inhabitants of Manhattan Island are not. That Baffin Islanders are "happier" than Manhattan Islanders still remains an unverifiable proposition.

[5]My account of Gordon's work is taken from Anastasi and Foley (1). I do not have access to the original source.

the Porteous Maze Test to Australian aborigines. Yerkes (24) reported a similar result with respect to different socio-economic groups, the increment of score with age being less for a low-status than for a high-status group. This finding, however, does not seem to have been clearly substantiated by later work.

Unequivocal interpretation of the above findings is not possible since the tests used at different age levels were somewhat different; but it seems reasonable to conclude that for any particular test the change in performance with age may vary markedly from one cultural group to another. In one the increment of test performance with age may be substantial, in another negligible. This must be mainly due to the demands of the cultural environment, which dictate what shall be learned and at what age.

If we accept this as the most plausible view, the inference may be drawn that those abilities that are of importance in a particular cultural environment, and that may be expected to correlate with performance in the important activities which the culture demands, are those which show a pronounced increment with age.[6] It is possible, that the abilities which the ordinary intelligence test defines are fairly independent of the types of activity which make for success in, say, a canal-boat culture, provided possibly that some minimal level is attained. It is possible that other abilities, if they could be defined, might show a marked increment with age, and be expected to correlate with the important classes of activity which make for survival in that culture.

The usual tests of intelligence correlate with a wide variety of classes of activity in our "more privileged" Western cultures. Such tests show substantial increase in performance with age until about the age of 17. We may speculate that tests which show a small age increment may correlate only with very specialized types of performance, and be of restricted usefulness.

If this line of inference could be substantiated, we would have available a criterion for the selection of tests which were likely to prove useful in a particular cultural environment. We would select those types of test material which had a high correlation with age and discard those

[6]This hypothesis assumes a direct relationship between the abilities that a particular form of education fosters in the child and the demands that the culture imposes at the adult level. In the history of education many cases of incongruity exist between education and the changing demands of a culture. Probably only in the most stable culture is a really high degree of congruence attained. Incongruity may enrich a culture, may destroy it, or may produce other effects. For example, if we impose our rather urbanized system of education, with its great emphasis on verbal abilities, on isolated Newfoundland communities, either these communities may be greatly enriched or extensive migration may occur. The latter alternative is more probable.

which had a low one. Age would then be a general criterion for the validation of tests correlating with performance on important classes of activity demanded by a particular culture.

Culture and Factor Theory

There are implications for factor analysis in the theory proposed in this paper. The following observations are speculative and call for more elaboration and refinement than can be afforded here.

The hypothesis of differential transfer which has been proposed implies that the factors or underlying parameters which transfer, and either facilitate or inhibit performance of a task, are not invariant with respect to the stage of learning at which the task is performed. Although a task may appear on superficial inspection to be the "same" at different stages of learning, operationally—in terms of factorial content—it may be different. Further, since the level of performance of a particular class of task may differ markedly from one culture to another, depending on the cultural dictum concerning what is learned and when, it follows that the factorial composition of tests may differ markedly from one culture to another. This simply means that, through learning, individuals in diverse cultures may bring different abilities to bear on the solution of an identical problem. Factorial invariance presumably applies only within the framework of a clearly defined cultural group, and has no broad cross-cultural implications. Were it technically feasible to construct a battery of tests which could be appropriately administered to a randon sample of adult subjects both in Toronto and in the Newfoundland outports, I have no doubt that we should find marked differences between the two groups in the factorial composition of many of the tests used. What we know about factors has reference to our own highly urbanized culture, which fosters the acquisition of certain verbal and reasoning abilities; our knowledge should not be presumed to extend beyond this.

The Problem of a General Factor

Spearman (**16**) strove to show that a general intellective factor operated in the performance of many mental tasks. Thurstone's (**19**) attempt to disprove Spearman's theory was not conclusive, since most of the factors in his classificatory system are correlated factors. The concept of a general factor is still with us. In the light of present knowledge this concept derives from the fact that many abilities, identified factorially in some loosely defined domain of intellectual activity, are not independent of one another. As currently regarded, it does not imply that all the innumerable identifiable abilities within that domain are correlated. Limitations on the generality of a general factor are not clearly prescribed.

A scrutiny of the general factor problem suggests that we must consider two aspects of it. We must account not only for the fact that many abilities are in some degree positively correlated with one another, but also for the fact that they are in some degree differentiated from one another. A theory which accounts for one aspect of the problem does not necessarily account for the other. Thus a theoretical explanation may account for abilities being correlated, but it may not, if pressed to its ultimate conclusion, account for their not being perfectly correlated, or there not being one ability only. The theory of human ability proposed in this paper can account very simply for the former aspect of the problem. To account for the latter aspect is more difficult.

Let us accept the general proposition that all learning, with the exception of some which occurs very early in life, occurs in a context of prior experience. This means that an individual will learn more readily activities which are facilitated by prior acquisitions, and will learn less readily those activities which are not facilitated or are perhaps inhibited by prior learning. Since in the adult many abilities are regarded as overlearned acquisitions, it follows that in the development of distinctive abilities those abilities will tend to develop which are facilitated and not inhibited by each other. It follows, therefore, that the positive correlation between abilities, which gives rise to the notion of a general factor, can be accounted for in the present theory by the operation of positive transfer. Although this is undoubtedly an oversimplification, it does provide a simple and plausible explanation.

Can we account for the fact that many abilities, while correlated, are none the less clearly differentiated? In approaching this problem we may observe that any attempt to explain learning and the formation of abilities by transfer alone leads to an obvious absurdity, since it cannot explain how early learning can occur at all. To escape this difficulty I propose a two-factor theory of learning. This theory states that much learning, excluding some very early learning, involves not only transfer components which are common to prior learning and the learning of a new task, but also components which are specific to the new task. In terms of the factorial model, this means that variation in performance at various stages of a task can be accounted for in part by the variation in prior acquisitions, and in part by specific abilities that emerge and are formed during the process of learning the task itself. I should anticipate that, as the learning of certain classes of tasks continues through a series of stages, the variance attributable to general transfer components may decrease, whereas the variance attributable to abilities specific to the task itself, or common only to the task, will increase. Further, it seems reasonable that variation in early learning can be accounted for much less by

transfer than by other processes, since there are fewer prior acquisitions to transfer. In the adult, learning may be accounted for largely in terms of transfer and to a much lesser extent in terms of other processes.

In sum, it seems to me that what happens is this. Some early learning must occur which is independent of prior learned acquisitions. As the individual grows, learning is facilitated more and more by prior acquisitions. It is probable that transfer effects become continuously and increasingly more important with age. In the learning of a particular task, transfer effects are probably greater at the earlier than at the later stages of learning. Thus, as the learning of a particular task continues, the ability to perform it becomes gradually differentiated from, although not necessarily independent of, other abilities which facilitate its differentiation. Learning is clearly a process by which the abilities of man are differentiated from one another, but the process of differentiation is aided and abetted by the abilities which the individual already possesses.

In conclusion, therefore, we may account for a component general to many abilities in terms of the operation of positive transfer, and for the differentiation of abilities in terms of the learning process itself, which, according to the theory presented here, operates in such a way as to facilitate differentiation.

Summary

In an attempt to draw together crudely within the same scheme the study of learning and the study of human ability I have advanced the following views. Different environments result in the overlearning of certain patterns of behaviour, which, because they are overlearned, become more or less invariant with respect to repetition or cessation. A crude limit of performance is reached. What is spoken of as an ability, in conventional psychological usage, has reference to performance at some crude limit of learning. This applies to the abilities of the Thurstone system and to whatever is subsumed under the term "intelligence." Differences in ability are the results of the complex interaction of the biological propensities of the organism, prior learning, and the age at which prior learning occurs. The role of human ability in subsequent learning, for example, intelligence in relation to scholastic performance, can be viewed as a problem in transfer; the question is in what way prior overlearned acquisitions—the abilities—affect subsequent learning. Abilities may transfer differentially in any learning situation; that is, the abilities which transfer and produce their effects at one stage of learning may be different from those which transfer and produce their effects at a later stage of the same task. An individual may possess the necessary ability to perform a task adequately, but may lack the ability to learn to perform the task

under particular learning conditions. The investigation of problems emerging from these lines of argument involves an individual difference approach and use of the methods of factorial analysis. The implications of the theory for problems of the role of cultural factors in human ability are elaborated. The line of theory developed leads to the inference that those abilities which are of importance in a particular culture, and which may be expected to correlate with performance in the important activities demanded for survival in the culture, are those which show a pronounced increment with age. This provides a basis for test validation of a certain type. The implications of the theory of human ability for factor theory are considered. The inference is drawn that ostensibly similar tests may have different factorial compositions in different cultures and in different strata of the same culture. The problem of a general factor is examined. A two-factor theory of learning is proposed. The correlation among abilities is explained in terms of positive transfer, and their differentiation by the development of abilities specific to particular learning situations.

REFERENCES

1. Anastasi, Anne, & Foley, J. P. *Differential psychology.* New York: Macmillan, 1949.
2. Anderson, H. D., & Eells, W. C. *Alaska natives: a survey of their sociological and educational status.* Stanford, Calif.: Stanford University Press, 1935.
3. Asher, E. J. The inadequacy of current intelligence tests for testing Kentucky Mountain children. *J. genetic Psychol.*, 1935, **46**, 480–486.
4. Cattell, R. B. A culture-free intelligence test. *J. educ. Psychol.*, 1940, **31**, 161–179.
5. Cattell, R. B. A culture-free intelligence test: Evaluation of cultural influences on test performance. *J. educ. Psychol.*, 1941, **32**, 81–100.
6. Cook, T. W. Repetition and learning: I. Stimulus and response. *Psychol. Rev.*, 1944, **51**, 25–36.
7. Eells, Kenneth, Davis, Allison, *et al. Intelligence and cultural differences.* Chicago: Univer. of Chicago Press, 1951.
8. Gomulicki, R. Bronislaw. The development and present status of the trace theory of memory. *British J. Psychol. Monogr. Suppl.*, 1953, **29**.
9. Gordon, H. *Mental and scholastic tests among retarded children.* Educ. Pamphlet No. 44. London: Board of Education, 1923.
10. Hebb, D. O. *The organization of behavior.* New York: Wiley, 1949.
11. Hirsch, N. D. M. An experimental study of East Kentucky Mountaineers. *Genet. Psychol. Monogr.*, 1928, **3**, 183–244.
12. McGeoch, John A. *The psychology of human learning.* New York: Longmans Green, 1946.
13. Porteous, S. D. *The psychology of a primitive people.* New York: Longmans Green, 1931.
14. Simrall, D. Intelligence and the ability to learn. *J. Psychol.*, 1947, **23**, 27–43.

15. SHERMAN, M., & KEY, C. B. The intelligence of isolated mountain children. *Child Develpm.*, 1932, 3, 279–290.
16. SPEARMAN, C. *The abilities of man: their nature and measurement.* London: Macmillan, 1927.
17. THOMSON, G. H. *The factorial analysis of human ability.* (2nd Ed.) London: Univer. of London Press, 1946.
18. THURSTONE, L. L. *Multiple factor analysis.* Chicago: Univer. of Chicago Press, 1947.
19. THURSTONE, L. L. *Primary mental abilities.* Chicago: Univer. of Chicago Press 1938.
20. TILTON, J. W. Intelligence test scores as indicative of ability to learn. *Educ. psychol. Measmt.*, 1949, 9, 291–296.
21. TURNBULL, W. W. Socio-economic status and predictive test scores. *Canad. J. Psychol.*, 1951, 5, 145–149.
22. WOODROW, H. Interrelations of measures of learning. *J. Psychol.*, 1940, 10, 49–73.
23. WOODROW, H. The ability to learn. *Psychol. Rev.*, 1946, 53, 147–158.
24. YERKES, ROBERT M., & ANDERSON, HELEN M. The importance of social status as indicated by the results of the point-scale method of measuring mental capacity. *J. educ. Psychol.*, 1915, 137–150.